स्वयं सीखू सीरीज

ए.आइ. सँ वीडियो बनाउ

पटकथा सँ प्रकाशन धरि सम्पूर्ण व्यवहारिक पाठ्यक्रम

चित्र, स्क्रीनशॉट, अभ्यास, परियोजना, तकनीकी विस्तार आ भाषा-निखार सहित

गजेन्द्र ठाकुर
सम्पादक — विदेह मैथिली ई-पत्रिका
www.videha.co.in | ISSN 2229-547X
प्रथम संस्करण — २०२६

प्रकाशकीय विवरण

© २०२६ गजेन्द्र ठाकुर। सर्वाधिकार सुरक्षित। लेखकक लिखित अनुमतिक बिना एहि पुस्तकक सम्पूर्ण वा महत्त्वपूर्ण अंशक व्यावसायिक पुनर्मुद्रण, विक्रय वा रूपान्तरण नहि कयल जाय। शिक्षण, समीक्षा आ अनुसंधानक उचित उद्धरण लेल स्रोत स्पष्ट लिखब आवश्यक अछि।

एहि पुस्तकमे प्रयुक्त औजारक नाम, लोगो आ अन्तरफलक सम्बन्धित कम्पनीक ट्रेडमार्क अछि। स्क्रीनशॉट सभ केवल चरण बुझाबय, समीक्षा करय आ डिजिटल साक्षरता सिखाबय लेल सीमित रूपमे देल गेल अछि। अन्तरफलक, योजना, मूल्य आ संस्करण समयक संग बदलि सकैत अछि; तेँ एहि पोथीमे बटन-नाम रटबाक बदला स्थायी कार्य-सिद्धान्त, रचनात्मक निर्णय, नैतिकता आ उत्पादन-प्रवाह पर बल देल गेल अछि।

लेखक वा प्रकाशक कोनो सदस्यता-मूल्य, क्रेडिट, उपलब्धता वा क्षेत्रीय सुविधा स्थायी रहबाक गारंटी नहि दैत छथि। प्रकाशन सँ पहिने सेवा-नियम, गोपनीयता नीति, कॉपीराइट, सहमति आ मञ्चक वर्तमान नियम स्वयं जाँचू।

पुस्तक-रूप: ए४, चित्रात्मक स्वयं-अध्ययन मार्गदर्शिका। भाषा: मैथिली। प्रथम संस्करण: २०२६।

ISBN: 978-93-5943-562-6

भूमिका

वीडियो बनाबय लेल पहिने कैमरा, प्रकाश, स्टूडियो, अभिनेता, सम्पादक आ भारी बजटक आवश्यकता बुझल जाइत छल। कृत्रिम बुद्धि एहि व्यवस्था केँ समाप्त नहि कयलक, मुदा छोट रचनाकारक हाथमे नूतन साधन अवश्य देलक। आब एकटा स्पष्ट विचार, सुगठित पटकथा, दू-चारि चित्र, सुचिन्तित प्रॉम्प्ट आ सामान्य सम्पादन-कौशलसँ शिक्षण-वीडियो, पुस्तक-परिचय, लोककथा, प्रचार-चित्र, लघु चलचित्र, गीत-दृश्य, वृत्तचित्रक सहायक दृश्य आ सामाजिक मञ्चक लघु सामग्री बनाओल जा सकैत अछि।

एहि पुस्तकक उद्देश्य कोनो एक कम्पनीक बटन याद करब नहि, बरु एहन मूल कौशल विकसित करब अछि जे औजार बदललाक बादो काम आए। एहि कारणेँ प्रत्येक पाठमे पहिने सिद्धान्त, तकर बाद स्क्रीनशॉट, फेर हाथक अभ्यास, स्व-जाँच आ एकटा छोट परियोजना देल गेल अछि।

मैथिली रचनाकारक सामने दूटा विशेष चुनौती अछि: पहिल, मातृभाषामे स्वाभाविक पटकथा आ स्वर; दोसर, स्थानीय संस्कृति केँ रूढ़ छवि वा कृत्रिम सजावटमे सीमित नहि करब। पुस्तक अहाँ केँ मिथिलाक परिवेश, वस्त्र, स्थापत्य, लोककला, उच्चारण आ सामाजिक संवेदनशीलताक उपयोग जिम्मेदारीसँ करबाक प्रशिक्षण दैत अछि।

ए.आइ. सहायक अछि, लेखक नहि। अन्तिम अर्थ, सत्यता, नैतिक जिम्मेदारी आ सौन्दर्य-बोध मनुष्यक हाथमे रहैत अछि। अहाँ जतेक स्पष्ट निर्णय लेब, परिणाम ततेक मौलिक होयत।

ई पुस्तक केना पढ़ी

पुस्तक तीन स्तर पर बनाओल गेल अछि। पहिल स्तरमे अहाँ बिना पूर्व अनुभवक मोबाइल वा ब्राउजरसँ छोट वीडियो बनायब सीखब। दोसर स्तरमे दृश्य-पट्ट, कैमरा-भाषा, प्रॉम्प्ट, स्वर आ सम्पादन पर नियन्त्रण बढ़ायब। तेसर स्तरमे लम्बा परियोजना, नैतिक प्रकाशन आ नियमित उत्पादन-प्रणाली विकसित करब।

1. प्रत्येक अध्यायक आरम्भमे देल उद्देश्य पढ़ू।
2. स्क्रीनशॉटमे अंकित लाल संख्या केँ पाठक चरणसँ मिलाउ।
3. प्रत्येक अभ्यासक फल अलग फोल्डरमे सुरक्षित राखू।
4. प्रॉम्प्टक पहिल परिणाम केँ अन्तिम नहि मानू; छोट परिवर्तन कऽ तीन-चारि रूप बनाउ।
5. अध्याय १५क तीन परियोजनामे कमसँ कम एकटा पूर्ण करू।
6. अन्तिम २१ दिनक कार्यक्रम पूरा कऽ अपन नमूना-सङ्ग्रह बनाउ।

विषय-विन्यास

- भाग १ — स्वयंसीखू मुख्य पाठ्यक्रम: मूलभूत अवधारणा सँ प्रकाशन, परियोजना आ २१-दिनक कार्यक्रम धरि
- भाग २ — विस्तारित व्याख्या आ अनुप्रयोग: योजना, स्टोरीबोर्ड, उपकरण, स्वर, अवतार, सङ्गीत, मैथिली सामग्री आ अभ्यास
- भाग ३ — उन्नत तकनीकी अध्ययन: २०२६ क प्रमुख वीडियो-जेनरेटर, प्रॉम्प्ट-रचना, कैमरा-प्रकाश, स्वर-अनुकरण, लिप-सिंक, VBench आ एपीआई कार्यप्रवाह
- परिशिष्ट — प्रॉम्प्ट-सङ्ग्रह, कार्यपत्रक, शब्दावली, उपकरण-दिशा; अन्तमे चयनित ग्रन्थ-सूची

भाग १
स्वयंसीखू मुख्य पाठ्यक्रम

ए.आइ. वीडियो की आ की नहि

एहि अध्यायक बाद अहाँ -

- जनरेटिव वीडियो आ सामान्य सम्पादनमे अन्तर बुझब
- ए.आइ.क सामर्थ्य आ सीमा चिन्हब
- सही परियोजना चुनब

मुख्य शब्द: जनरेटिव वीडियो · पाठसँ वीडियो · चित्रसँ वीडियो · मानवीय सम्पादन

ए.आइ. वीडियो एहन चल-चित्र सामग्री अछि जकर किछु वा सम्पूर्ण दृश्य, गति, स्वर, उपशीर्षक, काट-छाँट वा रूपान्तरण मशीन-अधिगम आधारित औजारक सहायता सँ बनाओल गेल हो। एकर अर्थ ई नहि जे सम्पूर्ण चलचित्र एक वाक्य लिखिते तैयार भऽ जाय। व्यवहारमे सफल वीडियो अनेक छोट निर्णयक परिणाम होइत अछि।

चारि प्रमुख प्रकार

पाठसँ वीडियो: अहाँ दृश्यक लिखित वर्णन दैत छी; औजार ओहि आधार पर छोट क्लिप बनबैत अछि।

चित्रसँ वीडियो: अहाँ प्रारम्भिक चित्र दैत छी; प्रॉम्प्ट मुख्यत: गति, कैमरा आ वातावरण बतबैत अछि।

वीडियोसँ वीडियो: पुरान क्लिपक शैली, पृष्ठभूमि, वस्त्र, प्रकाश वा गति बदलल जाइत अछि।

सहायक ए.आइ.: पटकथा, स्वत: उपशीर्षक, शोर-निवारण, स्वर, सङ्गीत, रंग-सुधार वा लघु अंश बनाबयमे उपयोग।

ए.आइ. की नीक करैत अछि

- अप्राप्य वा महँगा बी-रोलक कल्पनात्मक विकल्प बनबैत अछि।
- स्थिर चित्रमे हलचल, कैमरा गति आ वातावरण जोड़ैत अछि।
- एकहि विचारक अनेक दृश्य-रूप शीघ्र बनबैत अछि।
- पटकथा, शीर्षक, दृश्य-सूची आ सम्पादन-योजना बनाबयमे गति दैत अछि।
- स्वर, डबिङ, उपशीर्षक आ आकार-परिवर्तनक श्रम घटबैत अछि।

ए.आइ. कोन ठाम कमजोर अछि

- लम्बा दृश्यमे चेहरा, हाथ, वस्तु आ वस्त्रक निरन्तरता बिगड़ि सकैत अछि।
- लिखित अक्षर, लोगो, मानचित्र, गणना वा ऐतिहासिक तथ्य गलत हो सकैत अछि।
- एक दृश्यमे बहुत पात्र वा एक-साथ अनेक क्रिया भ्रम उत्पन्न करैत अछि।
- मैथिली उच्चारण, स्थानीय शब्द आ सांस्कृतिक सूक्ष्मता स्वत: शुद्ध नहि मानल जा सकैत अछि।
- यथार्थ देखाइत दृश्य सत्य हो, ई आवश्यक नहि।

मूल नियम ए.आइ. सँ दृश्य बनाउ, मुदा तथ्य स्वयं सत्यापित करू; ए.आइ. सँ स्वर बनाउ, मुदा उच्चारण स्वयं सुनू; ए.आइ. सँ पटकथा बनाउ, मुदा अर्थ आ शैली अपन राखू।

स्व-जाँच

7. पाठसँ वीडियो आ चित्रसँ वीडियोमे मूल अन्तर की अछि?
8. लम्बा कथात्मक वीडियो एके बेर जनाबयक बजाय छोट क्लिपमे किएक बनाबी?
9. यथार्थ देखाइत ए.आइ. दृश्यक सत्यता के जाँचत?

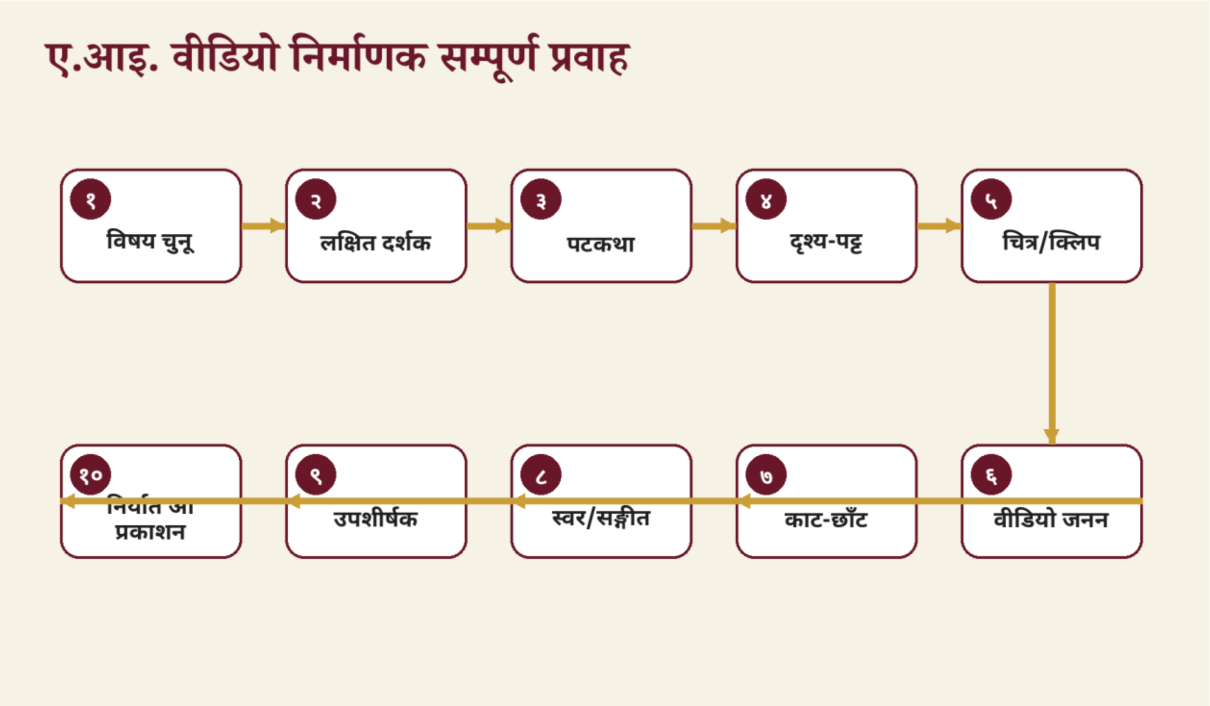
उत्तर-सङ्केत 1. पहिलमे दृश्य आ गति दुनू लिखल जाइत अछि; दोसरमे चित्र दृश्य निर्धारित करैत अछि आ प्रॉम्प्ट मुख्यतः गति बतबैत अछि। | 2. निरन्तरता, नियन्त्रण आ पुनःनिर्माण नीक रहैत अछि। | 3. मानवीय रचनाकार आ प्रकाशक।

विचारसँ प्रकाशन धरि कार्यप्रवाह

एहि अध्यायक बाद अहाँ -

- दस चरणक उत्पादन-प्रणाली अपनायब
- काजक क्रम बिगड़लासँ होमय वाला नुकसान बचायब
- प्रत्येक चरणक फल सुरक्षित राखब

मुख्य शब्द: कार्यप्रवाह · संस्करण · दृश्य-सूची · गुणवत्ता-जाँच



चित्र २.१ - ए.आइ. वीडियो निर्माणक दस चरण

मौलिक शिक्षण-चित्र

किएक क्रम आवश्यक अछि

नव सीखयवाला प्रायः पहिने औजार खोलि दृश्य जनाबय लगैत छथि। तखन विषय बदलैत अछि, अनुपात गलत होइत अछि, स्वरक अवधि दृश्यसँ नहि मिलैत अछि आ अन्तमे बहुत क्लिप व्यर्थ जाइत अछि। सही क्रम समय, क्रेडिट आ मानसिक श्रम बचबैत अछि।

विषय: एक वाक्यमे लिखू: वीडियो ककरा लेल आ की कहत?

दर्शक: आयु, भाषा, मञ्च आ पूर्वज्ञान निर्धारित करू।

पटकथा: पहिने स्वरक पाठ लिखू; फेर दृश्य जोड़ू।

दृश्य-पट्ट: प्रत्येक ३-८ सेकेँडक दृश्यक चित्रात्मक योजना बनाउ।

सामग्री: अपन फोटो, लोगो, मानचित्र, ध्वनि आ स्रोत व्यवस्थित करू।

जनन: पहिने कम गुणवत्ता वा छोट अवधिक परीक्षण बनाउ।

सम्पादन: सर्वोत्तम क्लिप चुनि क्रम, गति आ संक्रमण सुधारू।

स्वर/ध्वनि: स्वर स्पष्ट, सङ्गीत नीचाँ, प्रभाव सीमित राखू।

उपशीर्षक: मैथिली वर्तनी स्वयं जाँचू; स्वतः लिखल पाठ पर निर्भर नहि रहू।

निर्यात/प्रकाशन: सही अनुपात, रेजोल्यूशन, नाम, विवरण आ ए.आइ. प्रकटीकरण चुनू।

संस्करण-नामक साधारण नियम

एकहि फाइल केँ बेर-बेर “final” लिखि बचेनाइ भ्रम उत्पन्न करैत अछि। दिनाङ्क, परियोजना आ संस्करण जोड़ू: mithila_pokhari_v01_2026-08-04.mp4। सुधारक बाद v02, v03 लिखू। स्वीकृत फाइलमे FINAL जोड़ि सकैत छी, मुदा पुरान संस्करण नहि मिटाउ।

अभ्यास २.१ - एक मिनट वीडियो योजना

10. विषय चुनू: “मखान खेतसँ थारी धरि” वा अपन रुचिक विषय।
11. एक वाक्यमे दर्शक आ उद्देश्य लिखू।
12. छः दृश्यक क्रम लिखू।
13. प्रत्येक दृश्य लेल अवधि, चित्र आ स्वरक एक पंक्ति लिखू।

अभ्यासक फल एक पृष्ठक उत्पादन-योजना, जकर आधार पर अगिला अध्यायमे पटकथा बनत।

उपकरण, खाता आ संचिका-व्यवस्था

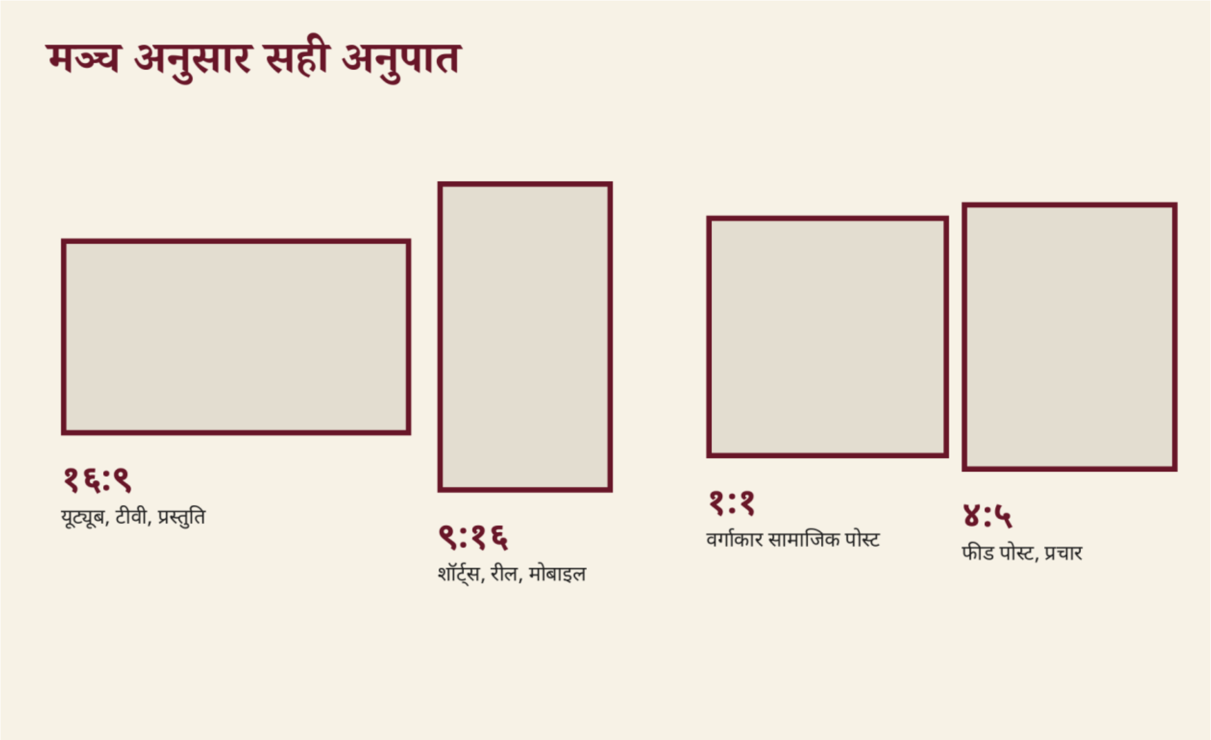
एहि अध्यायक बाद अहाँ -
• मोबाइल आ कम्प्यूटरक न्यूनतम तैयारी करब
• सही अनुपात आ रेजोल्यूशन चुनब
• सामग्रीक सुरक्षित फोल्डर बनायब
मुख्य शब्द: ब्राउजर · क्रेडिट · अनुपात · रेजोल्यूशन · बैकअप

न्यूनतम उपकरण

- स्थिर इंटरनेट; जनरेटिव वीडियो फाइल भारी होइत अछि।
- आधुनिक ब्राउजर; अनेक वेब औजार Chrome वा Safariक नवीन संस्करणक सलाह दैत अछि।
- कमसँ कम ५-१० जीबी खाली स्थान; असफल परीक्षण सेहो सुरक्षित राखल जा सकैत अछि।
- हेडफोन; छोट ध्वनि-दोष मोबाइलक स्पीकर पर नहि सुनाइत अछि।
- अपन स्वर रिकॉर्ड करबाक लेल शान्त कोठली आ सामान्य माइक्रोफोन।

फोल्डर संरचना

प्रत्येक परियोजनाक मूल फोल्डरक भीतर 01_script, 02_storyboard, 03_images, 04_generated_clips, 05_voice, 06_music_sfx, 07_edit, 08_export, 09_sources नामक फोल्डर बनाउ। अंकित नामसँ क्रम सदा स्पष्ट रहत।



चित्र ३.१ - मञ्च अनुसार सामान्य अनुपात

मौलिक शिक्षण-चित्र

अनुपात पहिने चुनू

१६: ९ क्षैतिज वीडियो यूट्यूब, टीवी आ व्याख्यान लेल उपयुक्त अछि। ९: १६ ठाढ़ वीडियो मोबाइल शॉर्ट्स आ रील लेल। १: १ वर्गाकार पोस्टक लेल; ४: ५ सामाजिक फीडमे अधिक स्क्रीन घेरैत अछि। एक अनुपातमे बनल दृश्य कें दोसर अनुपातमे काटलासँ मुख्य पात्र वा पाठ बाहर जा सकैत अछि।

क्रेडिट बचाबयक उपाय

- पहिने ५ सेकेँडक परीक्षण बनाउ।
- एक बेरमे एके परिवर्तन करू; प्रॉम्प्ट, अवधि, कैमरा सब एक-साथ नहि बदलू।
- प्रारम्भमे ७२०पी पर्याप्त; अन्तिम स्वीकृत दृश्ये उच्च गुणवत्ता बनाउ।
- पुन:निर्माण पहिने दोषक नाम लिखू: चेहरा, गति, प्रकाश, पृष्ठभूमि वा कैमरा।
- एकहि दृश्यक तीन उपयोगी संस्करण बनलापर रुकू; अनन्त विकल्प रचनात्मकता नहि, निर्णयहीनता अछि।

सुरक्षा गोपनीय पाण्डुलिपि, अप्रकाशित व्यक्तिगत फोटो, पहचान-पत्र, फोन नम्बर वा निजी वीडियो अपलोड करबाक पहिने औजारक गोपनीयता नीति पढ़ू।

विचार, शोध आ पटकथा

एहि अध्यायक बाद अहाँ -

- एक वाक्यक विचार कें वीडियो-पटकथामे बदलायब
- हुक, मुख्य भाग आ समापन लिखब
- दृश्य आ स्वरक अलग स्तम्भ बनायब

मुख्य शब्द: हुक · वाचन · बी-रोल · दृश्य-समय · स्रोत

एक वाक्यसँ आरम्भ करू

सफल विचारमे विषय, दर्शक आ फल तीनू रहैत अछि। उदाहरण: “विद्यालयक विद्यार्थी लेल दू मिनटमे मिथिलाक पञ्जी-प्रथाक मूल अवधारणा सरल दृश्यसँ बुझायब।” ई वाक्य बतबैत अछि जे वीडियो इतिहासक सम्पूर्ण शोध नहि, परिचयात्मक व्याख्या अछि।

तीन-खंड संरचना

आरम्भक हुक: पहिल ३-८ सेकेंडमे प्रश्न, आश्चर्य, दृश्य वा प्रतिज्ञा।

मुख्य भाग: एक समयमे एक विचार; प्रत्येक विचारक संग उपयुक्त दृश्य।

समापन: सार, उपयोगी अगिला कदम, स्रोत वा आह्वान।

नमूना - ६० सेकेंडक पटकथा

समय	स्वर	दृश्य	पाठ/ध्वनि
०-५ सेकेंड	“एकटा बीज, एकटा पोखरि आ पूरा ग्रामक अर्थव्यवस्था - ई अछि मखान।”	भोरमे पोखरि पर हवाई दृश्य	शीर्षक धीरे उभरय
५-१८	मखानक खेती पानिक भीतर होइत अछि...	किसान, पात, फूल, पानीक नजदीकी दृश्य	पानी आ पक्षीक हलुक ध्वनि
१८-३५	फल निकालि सुखाओल जाइत अछि...	हाथक कामक क्रम	तीन छोट लेबल
३५-५०	भुनबाक बाद बीज फूटि सफेद मखान बनैत अछि।	भुनाइ आ फूटल मखान	हलुक ताल
५०-६०	मखान केवल खाद्य नहि, मिथिलाक जल-संस्कृतिक अंश अछि।	थारी, परिवार, पोखरि	स्रोत आ समापन

पटकथा लिखबाक नियम

- जे वाक्य बोलब कठिन, ओ सुनबो कठिन; छोट वाक्य लिखू।
- चित्र जे स्पष्ट करैत अछि, स्वरमे फेर ओही बात नहि दोहराउ।
- अङ्क, नाम आ उच्चारणक अलग सूची बनाउ।
- प्रत्येक तथ्यक स्रोत नोट करू; अन्तिम विवरणमे स्रोत देब आसान होयत।
- मैथिली वाचन लेल विरामचिह्न, दीर्घ-ह्रस्व आ कठिन संयुक्ताक्षर स्वयं पढ़ि जाँचू।

अभ्यास ४.१ - अपन पटकथा

14. अध्याय २क विषय लिअ।
15. ६०-९० सेकेंडक वाचन १३०-१८० शब्दमे लिखू।
16. प्रत्येक वाक्यक सामने दृश्य लिखू।
17. मोबाइलमे पढ़ि समय मापू; निर्धारित समयसँ अधिक हो तँ शब्द घटाउ।

अभ्यासक फल समयबद्ध, बोलि कऽ जाँचल, दृश्य-सहित पटकथा।

दृश्य-पट्ट आ कैमरा-भाषा

एहि अध्यायक बाद अहाँ -

- पटकथा कें छोट दृश्यमे विभाजित करब
- शॉट आकार आ कैमरा गति चुनब
- निरन्तरता बनौने राखब

मुख्य शब्द: दृश्य-पट्ट · क्लोज-अप · वाइड शॉट · पैन · डॉली · निरन्तरता

चित्र ५.१ - छः-दृश्य दृश्य-पट्ट

मौलिक कार्यपत्रक; प्रिन्ट कऽ उपयोग कयल जा सकैत अछि

एक दृश्य - एक मुख्य क्रिया

छोट जनरेटिव क्लिपमे एके मुख्य क्रिया राखू। “बालिका दौड़ैत अछि, किताब खोलैत अछि, पक्षी उड़ैत अछि आ कैमरा चारू कात घुमैत अछि” जेकाँ निर्देश अधिकतर अस्थिर परिणाम देत। एकरा चारि दृश्यमे बाँटू।

सामान्य शॉट आकार

अत्यन्त नजदीकी: आँखि, हाथ, अक्षर, बनावट; भावना वा विवरण।

नजदीकी: चेहरा वा मुख्य वस्तु; दर्शकक ध्यान केन्द्रित।

मध्यम: कमरसँ ऊपर पात्र; संवाद आ क्रिया।

विस्तृत: पूरा पात्र आ परिवेश; स्थानक परिचय।

अत्यन्त विस्तृत: भू-दृश्य, ग्राम, भवन वा विशाल घटना।

कैमरा गति

पैन: कैमरा बाम-दहिन घूमैत अछि।

टिल्ट: ऊपर-नीच घूमैत अछि।

डॉली इन: कैमरा विषयक नजदीक जाइत अछि।

डॉली आउट: कैमरा पाछाँ हटैत अछि।

ट्रैकिङ: विषयक संग-संग चलैत अछि।

स्थिर: कैमरा नहि हिलैत; शिक्षा, उत्पाद आ संवादमे उपयोगी।

निरन्तरता-पत्र

एक पात्र अनेक दृश्यमे रहत तँ ओकर उमेर, चेहरा, केश, वस्त्र, आभूषण, भाव, प्रकाश आ स्थानक विवरण एक “पात्र-पत्र”मे लिखू। सम्भव हो तँ एके सन्दर्भ-चित्र उपयोग करू। हर दृश्यक प्रॉम्प्टमे अनावश्यक नूतन विशेषण नहि जोड़ू।

> दृश्य-संयम चलचित्रात्मकता कैमरा सदा घुमाबयसँ नहि अबैत अछि। कतेको प्रभावशाली दृश्य स्थिर कैमरा, स्पष्ट रचना आ सूक्ष्म मानवीय गतिसँ बनैत अछि।

प्रॉम्प्ट लिखबाक कला

> एहि अध्यायक बाद अहाँ -
> - सात अङ्गक प्रॉम्प्ट बनायब
> - पाठसँ वीडियो आ चित्रसँ वीडियो लेल अलग निर्देश लिखब
> - पुनरावृत्त सुधार करब
>
> मुख्य शब्द: प्रॉम्प्ट · निषेधात्मक निर्देश · कैमरा · शैली · बीज

चित्र ६.१ - सफल प्रॉम्प्टक सात अङ्ग

मौलिक शिक्षण-चित्र

प्रॉम्प्टक मूल सूत्र

शॉट प्रकार + मुख्य विषय + स्पष्ट क्रिया + स्थान/समय + दृश्य शैली + कैमरा गति + प्रकाश/भाव। एहि क्रमक उद्देश्य भाषा सजाबय नहि, निर्णय स्पष्ट करब अछि।

नमूना १ - पाठसँ वीडियो

औजारमे देबाक प्रॉम्प्ट A wide cinematic dawn shot of a Mithila village pond, lotus leaves and makhana plants moving gently in the breeze, two fishermen in the far background, slow forward dolly, soft golden light, natural colors, realistic documentary style, calm water reflections, no text, no logo.

मैथिली अर्थ: भोरक विस्तृत चलचित्रात्मक दृश्य; मिथिला ग्रामक पोखरि, कमल-पात आ मखानक पात हलुक हवा मे हिलैत, दूर दू मछुआर, कैमरा धीरे आगाँ बढ़ैत, कोमल सुनहला प्रकाश, स्वाभाविक रंग, वृत्तचित्र-जेकाँ यथार्थ शैली, शान्त प्रतिबिम्ब; कोनो पाठ वा लोगो नहि।

नमूना २ - चित्रसँ वीडियो

गति-कैँद्रित प्रॉम्प्ट The woman remains seated and looks toward the pond. Her sari edge moves gently in the breeze. Small ripples spread across the water. The camera makes a very slow push-in. Keep facial features, clothing, background architecture and colors unchanged.

चित्र पहिने विषय, वस्त्र आ स्थान निर्धारित कऽ चुकल अछि; एहि कारणेँ प्रॉम्प्टमे फेर चेहरा वर्णन करबाक बजाय गति आ जे बदलनाइ नहि चाही, से लिखल गेल अछि।

कमजोर प्रॉम्प्ट किएक असफल होइत अछि

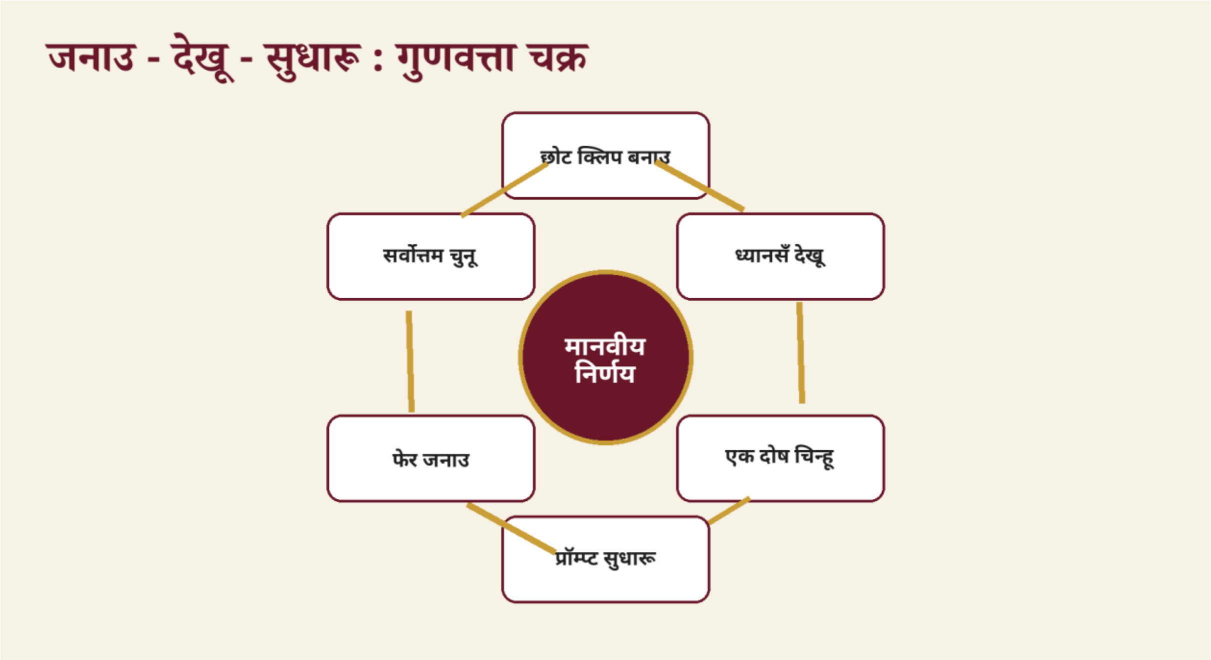
“सुन्दर वीडियो बनाउ”: सुन्दरक अर्थ अस्पष्ट; विषय, क्रिया आ कैमरा नहि।

बहुत विशेषण: “अद्भुत, दिव्य, महान, जादुई...” दृश्य निर्णय नहि दैत अछि।

अनेक क्रिया: छोट अवधिमे क्रम टूटैत अछि।

विरोधी निर्देश: “स्थिर कैमरा” आ “तेज घुमैत कैमरा” एकसाथ।

लिखित पाठ: जनरेटिव दृश्यमे अक्षर बिगड़ि सकैत अछि; पाठ बादमे सम्पादकमे जोड़ू।



चित्र ६.२ - जनाउ, देखू आ सुधारू चक्र

मौलिक शिक्षण-चित्र

एक बेरमे एक परिवर्तन

पहिल क्लिपमे कैमरा तेज अछि तँ केवल “very slow, almost static camera” जोड़ू। चेहरा, शैली, प्रकाश आ अवधि एकसाथ बदलला पर सुधारक कारण बुझब कठिन होयत। प्रत्येक परीक्षणक प्रॉम्प्ट सुरक्षित राखू।

अभ्यास ६.१ - तीन रूप

18. एके दृश्य लेल यथार्थ वृत्तचित्र शैलीक प्रॉम्प्ट लिखू।
19. ओही दृश्यक २डी चित्रकला रूप लिखू।
20. ओही दृश्यक केवल चित्रसँ वीडियो गति-प्रॉम्प्ट लिखू।
21. तीनूमे बदलल शब्द रेखांकित करू।

अभ्यासक फल एक विषयक तीन नियंत्रित प्रॉम्प्ट, जे शैली आ इनपुटक अन्तर स्पष्ट करैत हो।

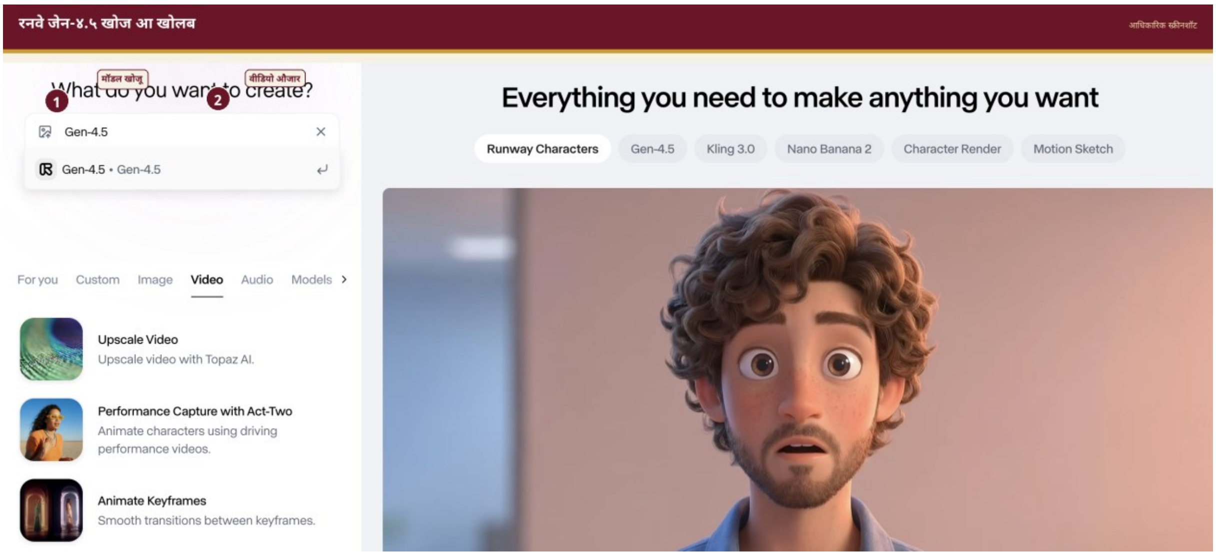
पाठसँ वीडियो: वेब औजार

एहि अध्यायक बाद अहाँ -

- रनवे जेन-४.५ जेकाँ वेब औजारमे क्लिप बनायब
- मॉडल, अनुपात, अवधि आ गति चुनब
- पुरान ट्यूटोरियलसँ बचब

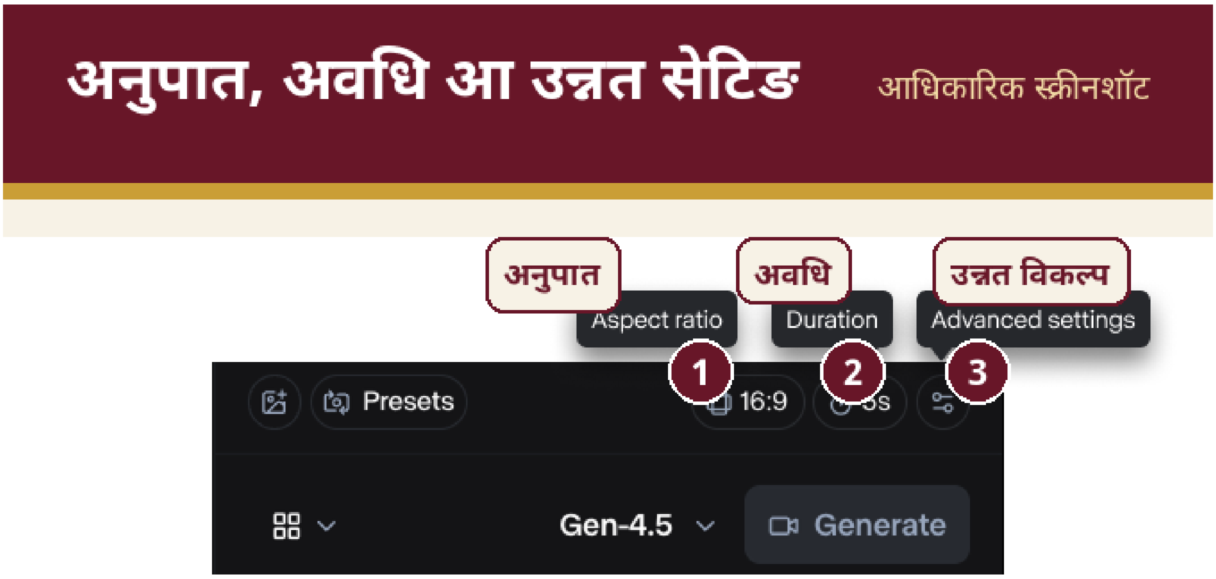
मुख्य शब्द: रनवे · फायरफ्लाइ · मॉडल · अवधि · एफपीएस

वेब-आधारित औजारक नाम आ रूप तेजी सँ बदलैत अछि, मुदा मूल प्रक्रिया लगभग एके अछि: मॉडल चुनू, प्रॉम्प्ट लिखू, अनुपात आ अवधि तय करू, परीक्षण बनाउ, दोष पहचानू, फेर सुधारू।



चित्र ७.१ - रनवेमे जेन-४.५ खोजब

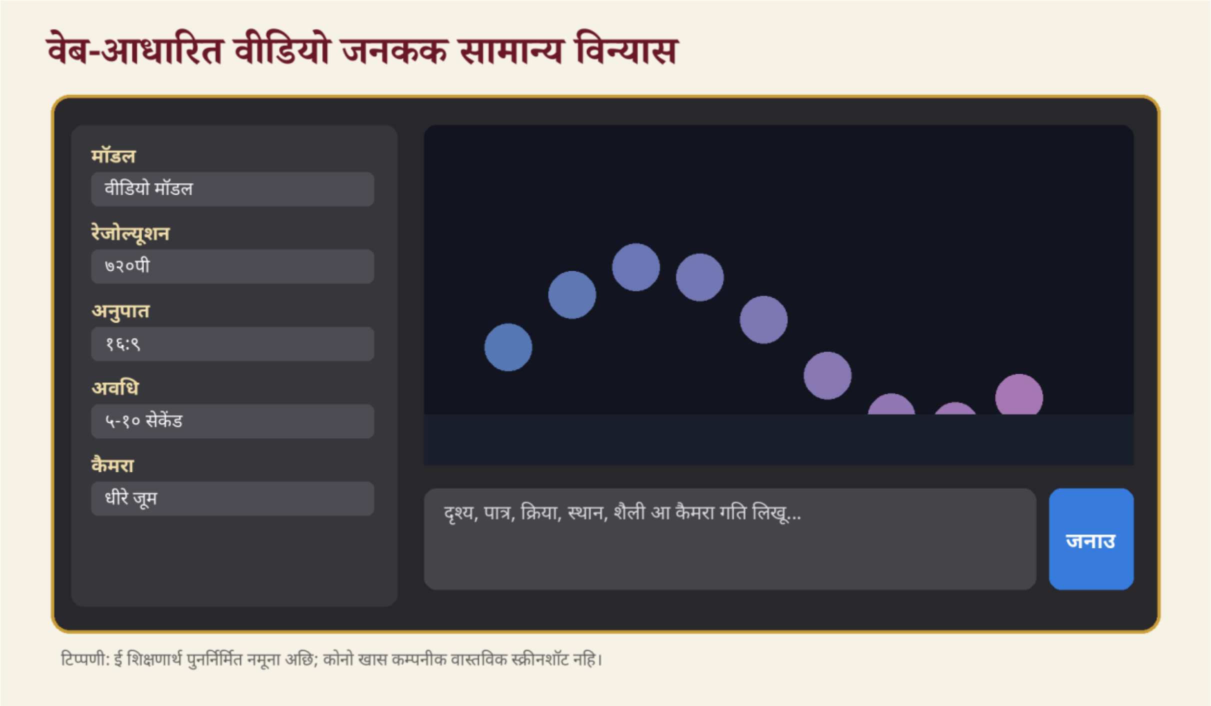
Runway Gen-4.5 जेहन आधुनिक जनरेटिव वीडियो औजार पाठसँ वीडियो आ चित्रसँ वीडियो दुनू कार्यप्रवाहक उदाहरण दैत अछि। पाठसँ वीडियोमे दृश्य आ गति दुनू स्पष्ट लिखू; चित्रसँ वीडियोमे मुख्य ध्यान गति, कैमरा आ निरन्तरता पर राखू। अवधि, रेजोल्यूशन आ अनुपातक विकल्प संस्करणानुसार बदलि सकैत अछि।



चित्र ७.२ - अनुपात, अवधि आ उन्नत विकल्प

चरण-दर-चरण

22. नव परियोजना खोलि Video वा सम्बन्धित जनरेटिव औजार चुनू।
23. मॉडल सूचीमे वर्तमान वीडियो मॉडल चुनू।
24. पाठसँ वीडियो लेल सात-अङ्ग प्रॉम्प्ट लिखू।
25. अनुपात लक्षित मञ्च अनुसार चुनू।
26. पहिल परीक्षण ५-८ सेकेंड राखू।
27. जनाउ; परिणाम पूरा देखू, केवल पहिल फ्रेम नहि।
28. सर्वोत्तम रूप डाउनलोड कऽ प्रॉम्प्टक संग सुरक्षित राखू।



चित्र ७.३ - वेब वीडियो जनकक सामान्य विन्यास

समझाबय लेल पुनर्निर्मित नमूना; वास्तविक कम्पनीक स्क्रीनशॉट नहि

Adobe Fireflyक सामान्य प्रवाह

आधुनिक जनरेटिव वीडियो औजारमे सामान्यतः मॉडल, रेजोल्यूशन, अनुपात, सन्दर्भ-चित्र वा सन्दर्भ-वीडियो, कैमरा विकल्प आ प्रॉम्प्ट चुनबाक सुविधा रहैत अछि। प्रभावी प्रॉम्प्टमे शॉट प्रकार, विषय वा पात्र, क्रिया, स्थान, प्रकाश आ दृश्य-शैली स्पष्ट क्रममे देनाइ उपयोगी अछि (Phoenix आ Taylor, २०२४; Bowen, २०२४)।

पुरान Sora ट्यूटोरियल OpenAIक Sora वेब आ ऐप अनुभव २६ अप्रैल २०२६ कैं बन्द भऽ गेल; Sora Videos API सेहो २४ सितम्बर २०२६ कैं बन्द होमयवाला अछि [९][१०]। अतः पुरान “Sora app” बटन आधारित पाठ पर समय वा धन नहि लगाउ।

स्व-जाँच

29. पाठसँ वीडियो प्रॉम्प्टमे की दू मुख्य वस्तु लिखब आवश्यक अछि?
30. पहिल परीक्षण छोट अवधि आ मध्यम रेजोल्यूशनमे किएक?
31. पुरान स्क्रीनशॉट देखला पर पहिल काज की?

उत्तर-सङ्केत 1. दृश्य आ गति। | 2. क्रेडिट बचत आ शीघ्र सुधार। | 3. आधिकारिक नवीन सहायता-पृष्ठ जाँचब।

चित्रसँ वीडियो: गति आ निरन्तरता

एहि अध्यायक बाद अहाँ -

- सही आरम्भिक चित्र चुनब
- गति-कैंद्रित प्रॉम्प्ट लिखब
- चेहरा आ पृष्ठभूमि स्थिर रखब

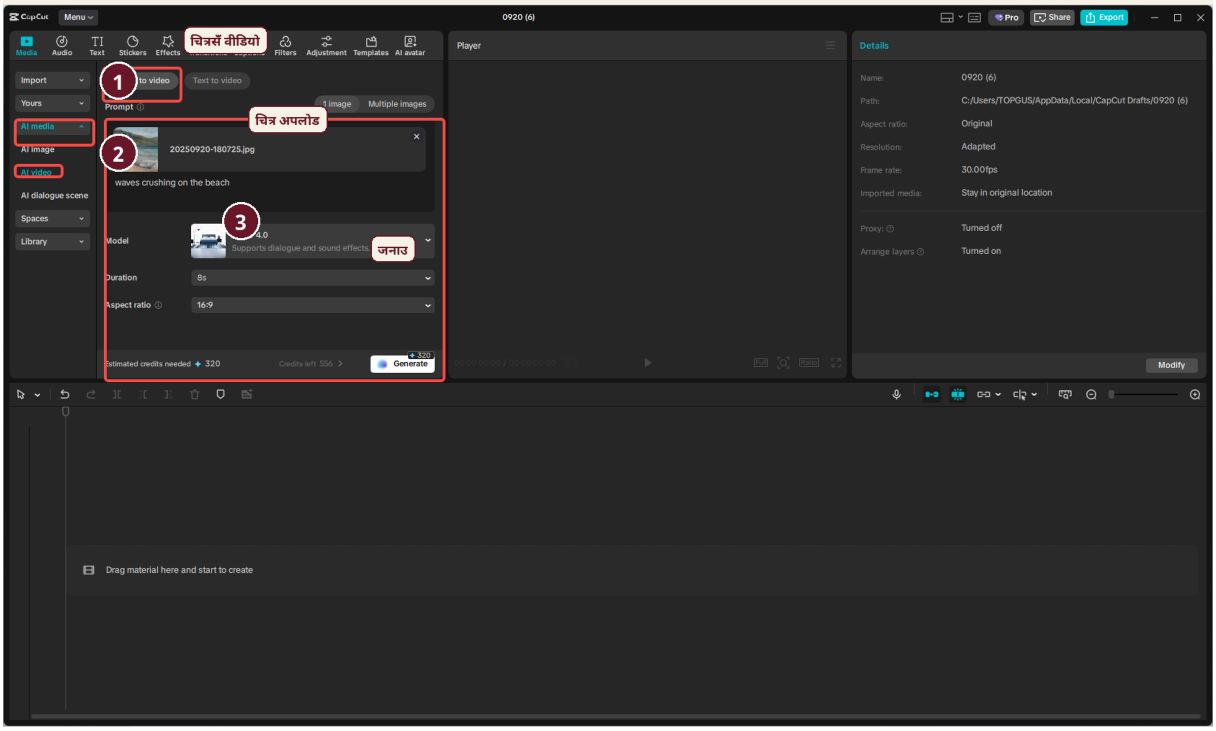
मुख्य शब्द: सन्दर्भ-चित्र · पहिल फ्रेम · गति · निरन्तरता · चित्राधिकार

चित्रसँ वीडियो नव सीखयवाला लेल प्रायः पाठसँ वीडियोसँ अधिक नियंत्रित अछि, कारण पात्र, वस्त्र, रंग आ रचना पहिने चित्रमे निश्चित रहैत अछि। मुदा खराब चित्रक दोष वीडियोमे बढ़ि सकैत अछि।

सही चित्रक गुण

- मुख्य विषय स्पष्ट आ पर्याप्त बड़ा हो।
- हाथ, चेहरा आ वस्तु कटायल नहि हो।
- प्रकाश एक दिशा सँ आ स्वाभाविक हो।
- पृष्ठभूमि अत्यधिक भीड़भाड़ वाला नहि हो।
- चित्रक अनुपात अन्तिम वीडियोसँ मिलैत हो।
- चित्र अपन हो, सार्वजनिक अधिकार वाला हो, वा उपयोगक लिखित अनुमति हो।

चित्रसँ वीडियो बनाबय वाला पट्टी आधिकारिक स्क्रीनशॉट



चित्र ८.१ - कॅपकटमे चित्रसँ वीडियो विकल्प

गति लिखबाक चारि स्तर

विषयक गति: आँखि झपकब, माथ घुमायब, चलब, हाथ उठायब।

परिवेशक गति: हवा, पानी, धुआँ, पात, बादल।

कैमरा गति: स्थिर, धीरे आगाँ, हलुक पैन।

स्थिरता निर्देश: चेहरा, वस्त्र, भवन, रंग आ मुख्य वस्तु नहि बदलू।

मिथिला चित्रकलासँ गति

मधुबनी चित्रकला वा रेखाचित्र कैं यथार्थ चलचित्र बनाबयक बजाय ओकर मूल सपाट रचना बचौने सूक्ष्म गति दियौ: माछक पूँछ हलुक हिलय, कमल-पात पर पानी चमकय, सीमा-रेखा स्थिर रहय, कैमरा अत्यन्त धीरे आगाँ बढ़य। “keep the original flat painted style and linework unchanged” जेकाँ निर्देश सहायक हो सकैत अछि।

अभ्यास ८.१ - एक चित्र, तीन गति

32. अपन अधिकार वाला एक स्वच्छ चित्र चुनू।
33. पहिल रूपमे केवल हवा वा पानीक सूक्ष्म गति लिखू।
34. दोसर रूपमे विषयक एक क्रिया जोड़ू।
35. तेसर रूपमे धीमा कैमरा गति जोड़ू।
36. तीनूमे चेहरा, वस्त्र आ पृष्ठभूमिक स्थिरता जाँचू।

अभ्यासक फल तीन पाँच-सेकेँड क्लिप आ तुलना-पत्र।

तत्काल पूर्ण वीडियो: कॅपकट कार्यप्रवाह

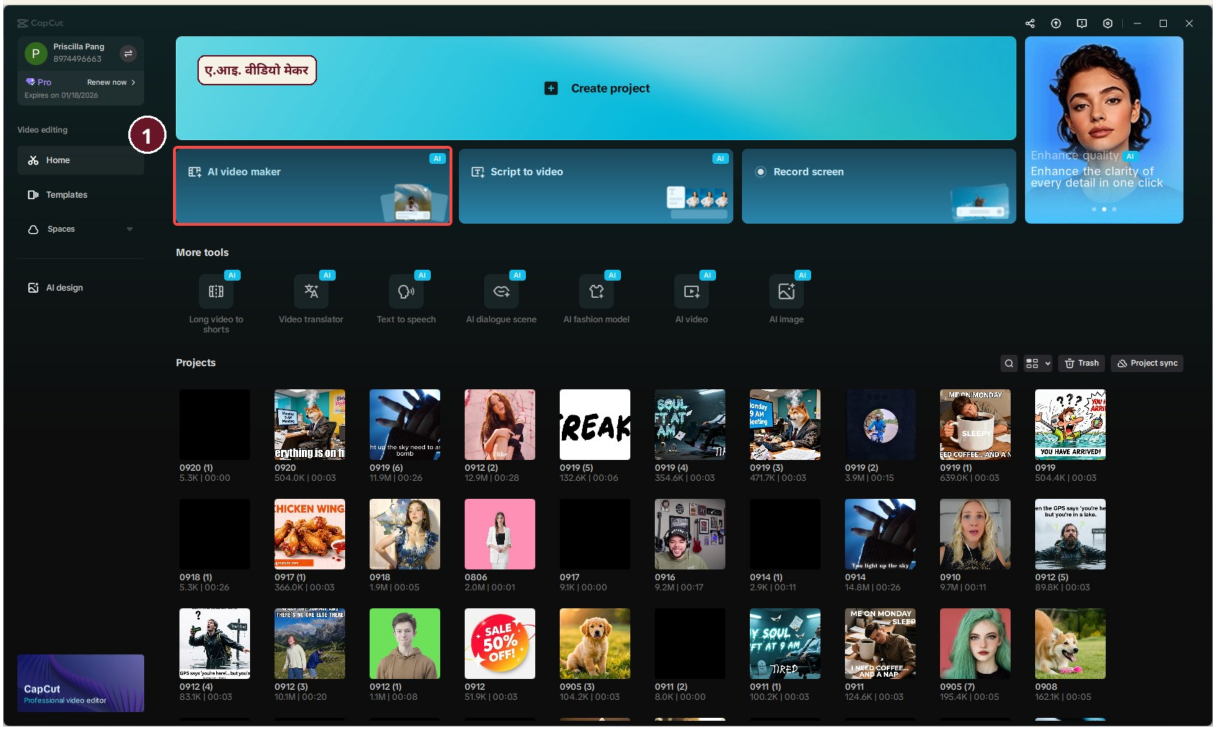
एहि अध्यायक बाद अहाँ -

- पटकथासँ स्वतः दृश्य बनायब
- दृश्य, पाठ आ सङ्गीत सुधारब
- सही सेटिङमे निर्यात करब

मुख्य शब्द: कॅपकट · ए.आइ. वीडियो मेकर · स्क्रिप्ट · टाइमलाइन · निर्यात

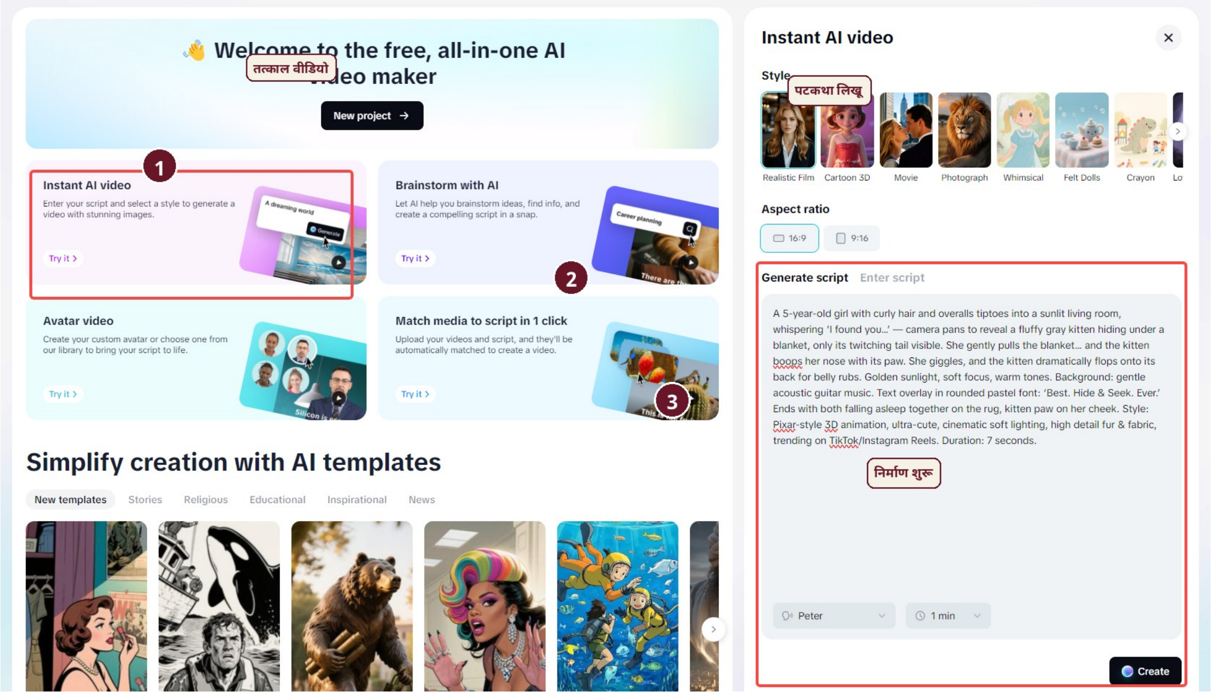
कॅपकट जेकाँ सम्पादक ए.आइ. जनन आ पारम्परिक सम्पादन कैं एक ठाम जोड़ैत अछि। एकरा कारणेँ नव सीखयवाला पटकथासँ प्रारम्भिक पूर्ण वीडियो बनाकऽ प्रत्येक दृश्य बादमे बदलि सकैत अछि। व्यवहारिक रूपसँ पूर्व-उत्पादन, सामग्री-निर्माण, टाइमलाइन सम्पादन, ध्वनि, उपशीर्षक आ निर्यातक क्रम अपनाएब उपयोगी अछि (Clark आदि, २०१९; Dancyger, २०२६)।

कॅपकटमे ए.आइ. वीडियो मेकर खोलब आधिकारिक स्क्रीनशॉट



चित्र १.१ - ए.आइ. वीडियो मेकर खोलू

विषय, शैली आ पटकथा भरब आधिकारिक स्क्रीनशॉट

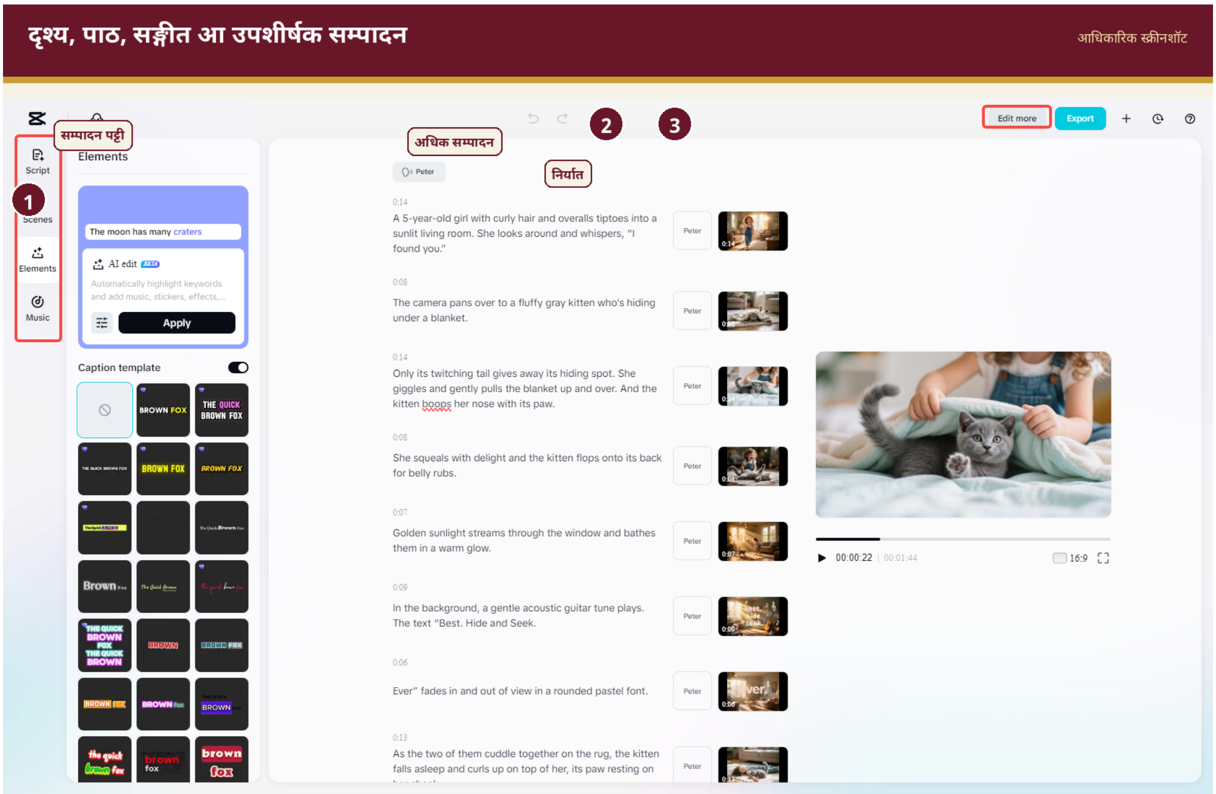


चित्र १.२ - शैली, अनुपात आ पटकथा भरू

तत्काल वीडियो बनाबयक क्रम

37. ए.आइ. वीडियो मेकर खोलू।
38. Instant AI video वा समान विकल्प चुनू।

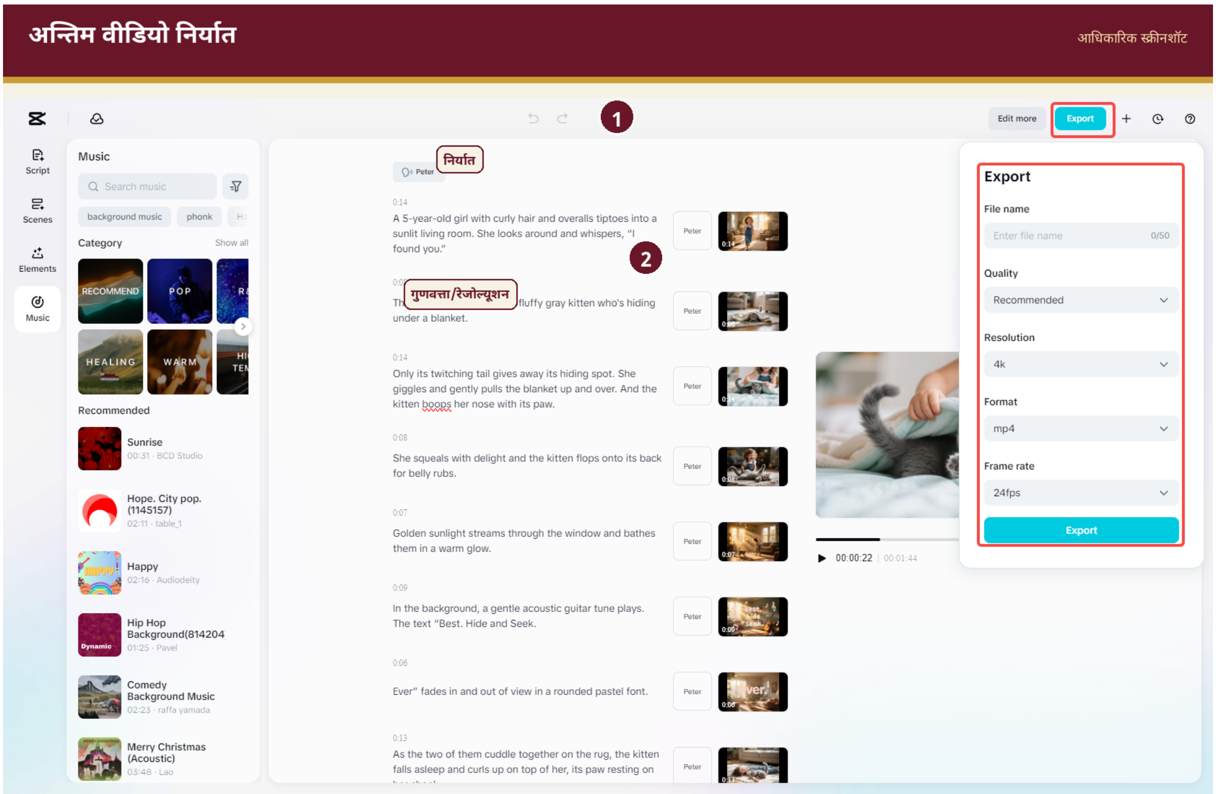
39. शैली चुनू; परियोजनाक विषयसँ मेल नहि खाइत “ट्रेंड” शैली नहि लिअ।
40. १६: ९ वा ९: १६ अनुपात चुनू।
41. पटकथा पेस्ट करू; दृश्य विभाजन पहिने स्वयं जाँचू।
42. प्रारम्भिक वीडियो बनाउ।
43. प्रत्येक दृश्यक चित्र, अवधि, पाठ, सङ्गीत आ क्रम सुधारू।



चित्र ९.३ - दृश्य, पाठ, तत्त्व आ सङ्गीतक सम्पादन

स्वतः चुनल दृश्य पर आँखि मूदि विश्वास नहि

पटकथामे “पञ्जी” शब्द रहला पर औजार सामान्य रजिस्टर, विवाह-दृश्य वा असम्बद्ध कार्यालय चुनि सकैत अछि। स्थानीय विषय लेल अपन चित्र, अभिलेखीय सामग्री, मानचित्र आ विशेष जनरेटिव क्लिप बदलि जोड़ू।



चित्र १.४ - अन्तिम निर्यात

निर्यातक सामान्य मान

उपयोग	अनुपात	रेजोल्यूशन	सुझाओ
यूट्यूब सामान्य	१६: ९	१९२०×१०८०	एमपी४, मूल फ्रेम दर
शॉर्ट्स/रील	९: १६	१०८०×१९२०	मुख्य पाठ सुरक्षित मध्य भागमे
अभिलेखीय प्रति	मूल	उच्चतम उपयोगी	कम संपीड़न, अलग सुरक्षित
पूर्वावलोकन	लक्षित	७२०पी	शीघ्र समीक्षा लेल

स्वर, सङ्गीत, ध्वनि आ डबिङ

एहि अध्यायक बाद अहाँ -

- स्वाभाविक मैथिली वाचन तैयार करब
- टीटीएस आ अपन रिकॉर्डिङमे चुनाव करब
- सङ्गीत आ ध्वनि संतुलित करब
- डबिङमे सहमति राखब

मुख्य शब्द: टीटीएस · स्वर-नकल · डबिङ · ध्वनि-प्रभाव · मिश्रण

चित्र १०.१ - स्वर आ ध्वनिक कार्यप्रवाह

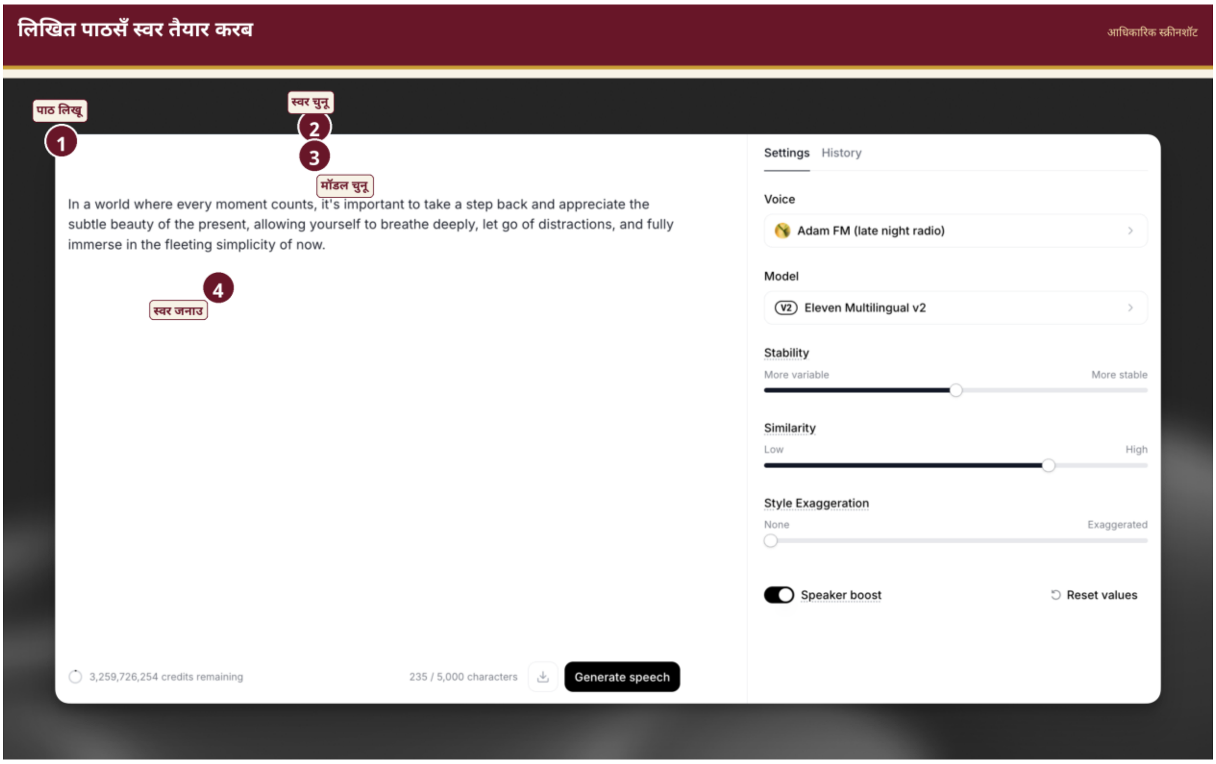
मौलिक शिक्षण-चित्र

तीन विकल्प

अपन स्वर रिकॉर्ड: मैथिली उच्चारण आ भाव पर सर्वाधिक नियन्त्रण; सामान्य माइक्रोफोन सेहो पर्याप्त, जँ कोठली शान्त हो।

पाठसँ स्वर: शीघ्र आ उपयोगी, मुदा भाषा-सूची, उच्चारण, लय आ योजना अनुसार सीमा रहि सकैत अछि।

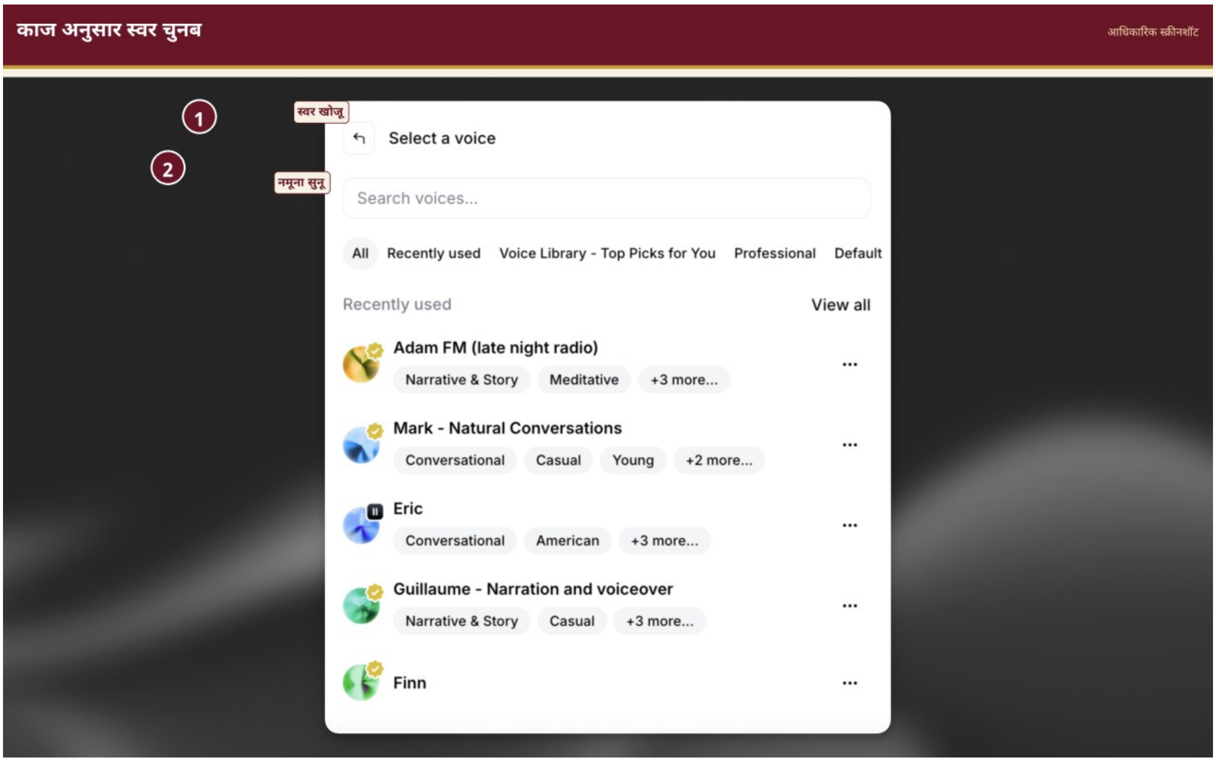
स्वर-परिवर्तन/नकल: केवल अपन स्वर वा स्पष्ट सहमति प्राप्त स्वर उपयोग करू; प्रतिरूपण नहि।



चित्र १०.२ - पाठ, स्वर, मॉडल आ जनन

पाठ-सँ-स्वर प्रणालीमे लिखल पाठ, उपयुक्त स्वर, भाषा वा उच्चारण आ जनन-सेटिङ चुनल जाइत अछि। मैथिली लेल पूरा पटकथा बनाबय पहिने २-३ वाक्यक परीक्षण करू; उच्चारण, दीर्घ-ह्रस्व, नाम, मुहावरा आ विराम स्वयं सुनि कऽ सुधारू। स्वर-निर्माण चाहे मानवीय रिकॉर्डिंग हो

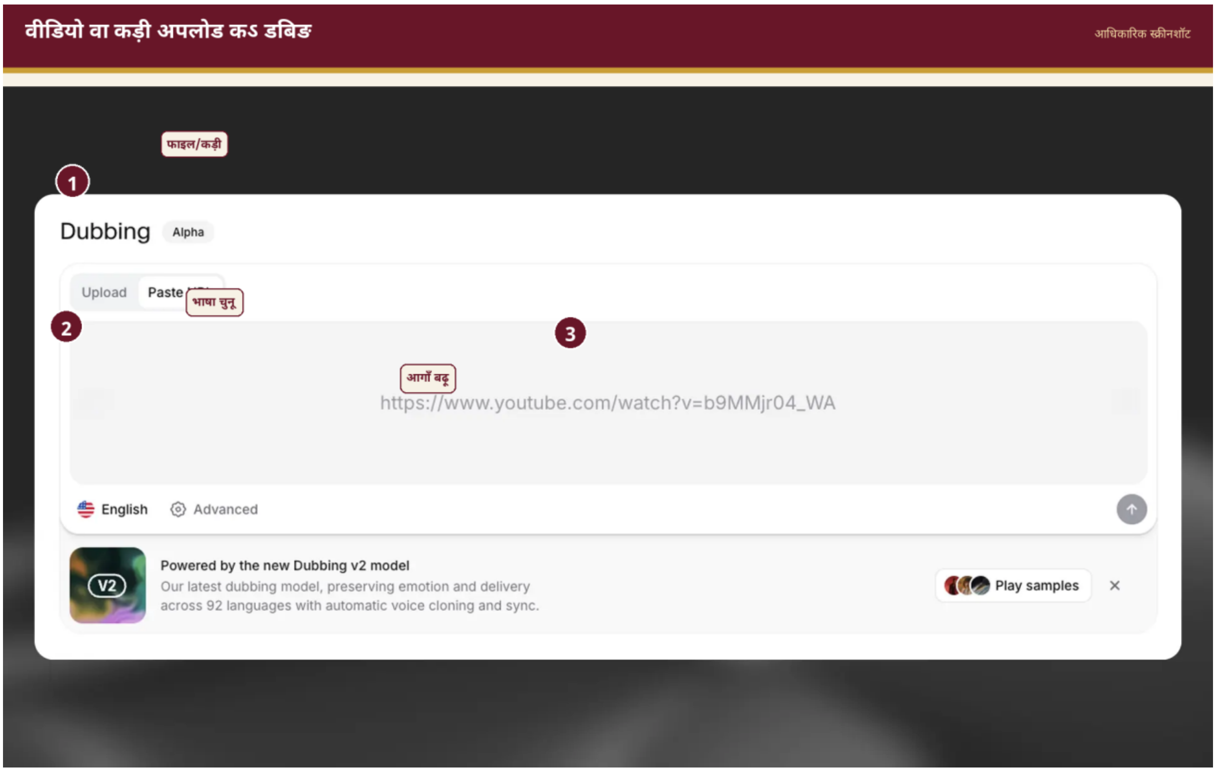
वा कृत्रिम, स्पष्टता, लय, ध्वनि-सन्तुलन आ अन्तिम मिश्रणक सिद्धान्त समान रूपेँ महत्त्वपूर्ण अछि (Rose, २०१५; Chattopadhyay, २०२४)।



चित्र १०.३ - काज अनुसार स्वर खोज आ नमूना सुनू

मैथिली उच्चारण सुधार

- कठिन नामक ध्वन्यात्मक रूप अलग लिखि परीक्षण करू।
- बहुत लम्बा अनुच्छेद नहि; २-४ वाक्यक खण्ड बनाउ।
- अल्पविराम, पूर्णविराम, दीर्घ विराम आ पंक्ति-विभाजनसँ लय नियंत्रित करू।
- विदेशी संक्षेप वा अङ्क केँ बोलबाक रूपमे लिखू।
- अन्तिम स्वर सुनि गलत शब्द हाथसँ बदलू; केवल तरंग-चित्र देखिकऽ स्वीकृत नहि करू।



चित्र १०.४ - फाइल वा कड़ीसँ डबिङ परियोजना

स्वचालित डबिङ कार्यप्रवाहमे मूल ऑडियो वा वीडियो, लक्ष्य भाषा, अनुवादित पाठ, स्वर-समानता, समय-साम्य आ अन्तिम मिश्रण शामिल होइत अछि। स्वचालित अनुवाद मैथिली अर्थ, सम्बोधन, लिंग, मुहावरा आ सांस्कृतिक सन्दर्भ बिगाड़ि सकैत अछि; तेँ डबिङक पाठ पहिने मानवीय सम्पादनसँ शुद्ध करू।

सङ्गीतक स्तर

वाचन मुख्य अछि। सङ्गीत एतेक नीचाँ राखू जे प्रत्येक व्यञ्जन स्पष्ट सुनाइ। भाव बदलबाक प्रत्येक ठाम सङ्गीत बदलनाइ आवश्यक नहि। कॉपीराइट-मुक्त वा उचित लाइसेंस वाला सङ्गीत उपयोग करू आ लाइसेंसक प्रति 09_sources फोल्डरमे राखू।

> सहमति ककरो स्वरक नकल करबाक पहिने लिखित अनुमति लिअ। मृत व्यक्ति, सार्वजनिक व्यक्ति वा परिवारक सदस्यक स्वर-सदृश प्रतिरूप सेहो भ्रम, अपमान वा कानूनी विवाद उत्पन्न कऽ सकैत अछि।

सम्पादन: टाइमलाइनसँ अन्तिम रूप

> एहि अध्यायक बाद अहाँ -
> - अप्रयुक्त क्लिप हटायब
> - लय, कट, संक्रमण आ रंग सुधारब
> - स्वर, सङ्गीत आ दृश्य समकालिक करब
>
> मुख्य शब्द: टाइमलाइन · कट · जे-कट · एल-कट · संक्रमण · रङ्ग-साम्य

सम्पादन चयनक कला अछि

ए.आइ. सँ दस क्लिप बनाओल गेल हो, मुदा अन्तिम वीडियोमे तीन श्रेष्ठ क्लिपे रहि सकैत अछि। जननमे लगल समयक कारण कमजोर क्लिप राखब “डूबल लागत”क भूल अछि। दर्शकक अनुभव महत्त्वपूर्ण अछि, जननक मेहनत नहि।

पहिल मोटा सम्पादन

44. पटकथाक क्रममे स्वर वा अस्थायी वाचन टाइमलाइन पर राखू।
45. प्रत्येक वाक्यक ऊपर सम्बन्धित दृश्य राखू।
46. दृश्यक आरम्भ-अन्तक अस्थिर फ्रेम काटू।
47. दोहराओल वा कमजोर दृश्य हटाउ।
48. सम्पूर्ण वीडियो एक बेर बिना रोकने देखू।

कट आ संक्रमण

साधारण सीधा कट अधिकांश स्थितिमे सर्वोत्तम अछि। फेड आरम्भ, समापन वा समय-परिवर्तनमे; डिसॉल्व स्मृति वा धीमा संक्रमणमे। प्रत्येक दृश्य पर अलग चमकदार संक्रमण पुस्तक-परिचय वा शैक्षणिक वीडियो केँ सस्ता देखाबि सकैत अछि।

जे-कट आ एल-कट

जे-कटमे अगिला दृश्यक ध्वनि पहिने सुनाइत अछि, तकर बाद चित्र बदलैत अछि। एल-कटमे पिछला दृश्यक स्वर अगिला चित्र पर किछु समय जारी रहैत अछि। ई तकनीक संवाद, साक्षात्कार आ वृत्तचित्रमे प्रवाह स्वाभाविक बनबैत अछि।

रङ्ग-साम्य

- सभ क्लिपक उजास लगभग समान करू।
- एक दृश्य अत्यधिक नीला, दोसर अत्यधिक पीयर हो तँ तापमान संतुलित करू।
- त्वचा-रंग अस्वाभाविक नहि हो।
- काला भागमे विवरण बचाउ; उज्जर भाग पूर्ण सफेद नहि जळय।
- ए.आइ. दृश्य पर अत्यधिक तीखापन नहि जोड़ू।

अभ्यास ११.१ - तीस सेकैंडक संपादन

49. पाँच क्लिप आ एक स्वर-फाइल टाइमलाइन पर राखू।
50. कमसँ कम एक क्लिप पूरा हटाउ।
51. तीन ठाम सीधा कट, एक ठाम फेड उपयोग करू।
52. सङ्गीत एतेक घटाउ जे स्वर स्पष्ट रहय।
53. मोबाइल आ हेडफोन दुनूमे देखू-सुनू।

अभ्यासक फल ३० सेकैंडक स्वच्छ, निर्यात योग्य क्रम।

उपशीर्षक, पहुँचयोग्यता आ मैथिली लिपि

एहि अध्यायक बाद अहाँ -
- स्वतः उपशीर्षक केँ शुद्ध करब
- देवनागरी फॉन्ट आ पंक्ति-विभाजन चुनब
- बिना ध्वनि देखनिहार दर्शक लेल वीडियो उपयोगी बनायब

मुख्य शब्द: उपशीर्षक · एसआरटी · देवनागरी · सुरक्षित क्षेत्र · वैकल्पिक पाठ

उपशीर्षक केवल सजावट नहि

बहुत दर्शक मोबाइल पर ध्वनि बन्द कऽ वीडियो देखैत छथि। श्रवण-बाधित दर्शक, शोर वाला स्थान, भाषा सीखयवाला छात्र आ कठिन नाम बुझयवाला व्यक्ति सभ लेल उपशीर्षक आवश्यक अछि।

मैथिली उपशीर्षकक नियम

- एक बेरमे अधिकतम दू पंक्ति।
- एक पंक्ति लगभग ३०-४२ अक्षर; मोबाइल पर परीक्षण करू।
- वाक्यक अर्थपूर्ण ठाम पंक्ति तोड़ू, शब्दक बीच नहि।
- उज्जर पाठक संग गाढ़ पृष्ठभूमि वा छाया; रंग मात्र पर निर्भर नहि।
- Noto Sans Devanagari जेकाँ स्पष्ट यूनिकोड फॉन्ट उपयोग करू।
- स्वत: लिखल नाम, अनुस्वार, चन्द्रबिन्दु, संयुक्ताक्षर आ मैथिली क्रिया स्वयं जाँचू।

पाठ कहाँ राखी

९: १६ वीडियोमे मञ्चक बटन दहिना आ नीचाँ रहैत अछि। मुख्य उपशीर्षक एकदम तल नहि; “सुरक्षित मध्य क्षेत्र”मे राखू। चेहरा वा मुख्य वस्तुक ऊपर पाठ नहि चढ़ाउ।

एसआरटी आ जळायल उपशीर्षक

एसआरटी: अलग पाठ-संचिका; दर्शक चालू-बन्द कऽ सकैत अछि, खोज आ अनुवादमे सहायक।

जळायल उपशीर्षक: चित्रक भीतर स्थायी; प्रत्येक मञ्च पर देखाइत, मुदा बादमे सुधार कठिन।

गुणवत्ता-जाँच उपशीर्षकक शुद्धता केवल पढ़ि नहि, वीडियो चलाकऽ जाँचू। पाठक समय स्वरसँ मिलैत अछि वा नहि, ई समान रूपसँ महत्त्वपूर्ण अछि।

यूट्यूब, शॉर्ट्स आ सामाजिक प्रकाशन

एहि अध्यायक बाद अहाँ -
- सही शीर्षक, थम्बनेल आ विवरण लिखब
- ए.आइ. सामग्रीक प्रकटीकरण करब
- एक वीडियो केँ अनेक मञ्च लेल रूपान्तरित करब

मुख्य शब्द: थम्बनेल · विवरण · प्रकटीकरण · मेटाडाटा · सीश्पीए

प्रकाशन पहिने अन्तिम पैकेज

- अन्तिम एमपी४
- थम्बनेल
- शीर्षकक तीन विकल्प
- दू अनुच्छेदक विवरण
- स्रोत आ लाइसेंस सूची
- उपशीर्षक फाइल
- ए.आइ. उपयोगक प्रकटीकरण

- सम्बन्धित कड़ी वा अगिला पाठ

शीर्षक

शीर्षक स्पष्ट प्रतिज्ञा करय, छल नहि। “३ मिनटमे ए.आइ. सँ मैथिली पुस्तक-परिचय वीडियो” स्पष्ट अछि। “ई रहस्य देखिकऽ चौंकि जाएब” जेकाँ भ्रामक शीर्षक अल्पकालिक क्लिक दे सकैत अछि, विश्वास नहि।

थम्बनेल

- एक मुख्य चेहरा वा वस्तु।
- ३-६ शब्दक बड़ा पाठ; वीडियो जनकमे नहि, डिजाइन औजारमे लिखू।
- मोबाइल आकारमे परीक्षण।
- वीडियोमे नहि रहल घटना वा व्यक्ति नहि देखाउ।
- ब्राण्ड-रंग सीमित आ एकरूप।

ए.आइ. प्रकटीकरण

यथार्थ-सदृश कृत्रिम वा परिवर्तित दृश्य आ स्वर दर्शककेँ भ्रमित कऽ सकैत अछि। तेँ वास्तविक व्यक्ति, घटना वा स्थानक कृत्रिम पुनर्निर्माण वा परिवर्तन स्पष्ट रूपेँ घोषित करू; शीर्षक, विवरण वा आरम्भिक सूचनामे बताउ जे कोन अंश कृत्रिम अछि। सामान्य रंग-सुधार, शोर-निवारण वा स्पष्ट काल्पनिक एनीमेशन अलग प्रकृतिक हो सकैत अछि; मूल सिद्धान्त पारदर्शिता, सहमति आ उत्तरदायित्व अछि (Coeckelbergh, २०२०; Dubber आदि, २०२०)।

> साफ विवरणक नमूना “एहि वीडियोमे किछु दृश्य कृत्रिम बुद्धि आधारित औजारसँ बनाओल गेल अछि। पटकथा, तथ्य-जाँच, मैथिली वाचन आ अन्तिम सम्पादन मानवीय रूपसँ कयल गेल अछि।”

एक सामग्री, तीन रूप

रूप	अवधि	अनुपात	काज
मुख्य वीडियो	३-८ मिनट	१६: ९	सम्पूर्ण कथा/पाठ
शॉर्ट	३०-६० सेकैंड	९: १६	एक प्रश्न वा मुख्य तथ्य
टिजर	१०-२० सेकैंड	९: १६/१: १	मुख्य वीडियो दिस नेब

कॉपीराइट, सहमति, डीपफेक आ सुरक्षा

> एहि अध्यायक बाद अहाँ -
> - अधिकारयुक्त सामग्री चुनब
> - चेहरा आ स्वरक सहमति बुझब
> - भ्रामक यथार्थ दृश्यसँ बचब
> - भारतीय नियमक दिशा बुझब
>
> मुख्य शब्द: कॉपीराइट · लाइसेंस · सहमति · डीपफेक · सिंथेटिक सूचना

चित्र १४.१ - नैतिक-सुरक्षा संकेत

मौलिक शिक्षण-चित्र

चारि प्रश्न

54. ई चित्र, वीडियो, सङ्गीत वा स्वर केकर अछि?
55. हमरा उपयोग, परिवर्तन आ प्रकाशनक अनुमति अछि?
56. दर्शक एकरा वास्तविक मानि भ्रमित हो सकैत अछि?
57. एहि सामग्रीसँ कोनो व्यक्ति, समुदाय वा संस्था कैं अनुचित हानि होयत?

अधिकारक प्रमाण राखू

प्रत्येक बाहरी सामग्रीक कड़ी, लेखक, लाइसेंस, डाउनलोड दिनाङ्क आ अनुमति 09_sources फोल्डरमे राखू। “इन्टरनेट पर भेटल” उपयोग-अधिकार नहि होइत अछि। ए.आइ. औजारक सेवा-नियम सेहो अलग-अलग होइत अछि; व्यावसायिक उपयोगक पहिने पढ़ू।

चेहरा आ स्वर

अपन फोटो वा स्वर उपयोग सरल अछि, मुदा आन व्यक्तिक चेहरा, निजी फोटो वा स्वर लेल स्पष्ट, सूचित आ लिखित सहमति आवश्यक मानू। अनुमति केवल “फोटो उपयोग”क हो, तँ ए.आइ. रूपान्तरण वा स्वर-नकल स्वतः शामिल नहि। बालकक सामग्री लेल अभिभावकीय अनुमति आ विशेष सावधानी आवश्यक।

भारतमे सिंथेटिक सामग्री

कृत्रिम वा परिवर्तित यथार्थ सामग्री प्रकाशित करैत समय सहमति, गोपनीयता, मानहानि, कॉपीराइट, व्यक्तित्व-अधिकार आ भ्रामक प्रस्तुतीकरणक जोखिम पर विचार करू। देश, मञ्च आ प्रसंगानुसार नियम भिन्न भऽ सकैत अछि; स्पष्ट लेबल, स्रोत-सत्यापन आ जिम्मेदार प्रकाशनक नीति अपनाउ, आ विवादास्पद वा व्यावसायिक उपयोगमे योग्य विधिक सलाह लिअ (Dubber आदि, २०२०; Liao, २०२०)।

की नहि करू

- ककरो एहन कथन करैत देखाउ जे ओ कहने नहि छथि।

- आपदा, चुनाव, युद्ध, बीमारी वा वित्त सम्बन्धी झूठा यथार्थ दृश्य बनाउ।
- आनक स्वर नकल कऽ फोन, विज्ञापन, समर्थन वा मजाक बनाउ।
- अभिलेखीय चित्रक तिथि/स्थान बदैल सत्य जेकाँ प्रस्तुत करू।
- ए.आइ. त्रुटि देखलाक बादो "औजार बनौलक" कहि जिम्मेदारी टारू।

कानूनी सीमा ई अध्याय सामान्य डिजिटल-साक्षरता मार्गदर्शन अछि, व्यक्तिगत कानूनी परामर्श नहि। कानून आ मञ्च-नियम बदलैत रहैत अछि।

तीन पूर्ण व्यवहारिक परियोजना

एहि अध्यायक बाद अहाँ -

- आरम्भसँ अन्त धरि सम्पूर्ण प्रक्रिया लागू करब
- भिन्न अवधि आ मञ्च लेल रूप बनायब
- मौलिक मैथिली सामग्री तैयार करब

मुख्य शब्द: लघु रील · शैक्षणिक वीडियो · पुस्तक-परिचय · उत्पादन-पत्र

परियोजना १ - ३० सेकैंडक दृश्य-कविता: "पोखरिक भोर"

उद्देश्य: ९: १६ मोबाइल रील, पाँच दृश्य, अपन वा अधिकारयुक्त चित्र, मैथिली काव्यात्मक वाचन।

- दृश्य 1: भोरक पहिल उजास पोखरिक पानि पर उतरैत अछि।
- दृश्य 2: मखानक पात बीच हवा अपन सूक्ष्म अक्षर लिखैत अछि।
- दृश्य 3: दूर नावक एक चोट, पानि मे गोल-गोल समय बनबैत अछि।
- दृश्य 4: ग्राम जागल नहि अछि, मुदा पक्षी सभ दिनक पहिल समाचार सुना रहल अछि।
- दृश्य 5: ई पोखरि जल नहि मात्र - स्मृति, जीविका आ जीवनक दर्पण अछि।

दृश्य प्रॉम्प्ट Vertical 9:16, a calm Mithila village pond at dawn, makhana leaves, soft mist, subtle ripples, distant wooden boat, slow cinematic push-in, natural documentary colors, peaceful mood, no text, no logo.

निर्माण क्रम

58. ९: १६ अनुपात चुनू।
59. एक विस्तृत, दू नजदीकी, एक नाव आ एक समापन दृश्य बनाउ।
60. अपन स्वर शान्त कोठलीमे रिकॉर्ड करू।
61. हलुक प्राकृतिक ध्वनि जोड़ू।
62. दू पंक्तिसँ अधिक उपशीर्षक नहि।
63. ए.आइ. दृश्यक प्रकटीकरण विवरणमे जोड़ू।

अभ्यासक फल ३० सेकैंडक एमपी४, थम्बनेल आ एसआरटी।

परियोजना २ - ३ मिनटक शैक्षणिक वीडियो: "मैथिली पुस्तक केना पढ़ी"

उद्देश्य: १६: ९ व्याख्यात्मक वीडियो, स्क्रीन रिकॉर्डिङ, पुस्तक-पृष्ठ, सरल एनिमेटेड सङ्केत आ स्पष्ट उपशीर्षक।

समय	विषय	दृश्य

०: ००-०: १५	हुक: पुस्तक खोललासँ पहिने तीन बात जाँचू।	पुस्तक, नोट, सरल आइकन, वक्ता वा स्क्रीन
०: १५-०: ५०	लेखक, विधा आ प्रकाशन-सन्दर्भ।	पुस्तक, नोट, सरल आइकन, वक्ता वा स्क्रीन
०: ५०-१: ३०	पहिल पठन: कथा वा तर्क बुझू।	पुस्तक, नोट, सरल आइकन, वक्ता वा स्क्रीन
१: ३०-२: १०	दोसर पठन: भाषा, प्रतीक, संरचना नोट करू।	पुस्तक, नोट, सरल आइकन, वक्ता वा स्क्रीन
२: १०-२: ४०	प्रश्न आ उद्धरणक नोट।	पुस्तक, नोट, सरल आइकन, वक्ता वा स्क्रीन
२: ४०-३: ००	सार आ अगिला पुस्तक।	पुस्तक, नोट, सरल आइकन, वक्ता वा स्क्रीन

दृश्य-सिद्धान्त शैक्षणिक वीडियोमे हर वाक्य लेल नूतन ए.आइ. दृश्य आवश्यक नहि। वास्तविक पुस्तक-पृष्ठ, स्क्रीन रिकॉर्डिङ, सरल रेखा, शीर्षक-पट्ट आ सीमित बी-रोल अधिक विश्वसनीय होइत अछि।

परियोजना ३ - ५ मिनटक पुस्तक-परिचय

उद्देश्य: पुस्तकक कथा खोलि नहि देत, मुदा विषय, शैली, ऐतिहासिक महत्त्व, लक्षित पाठक आ पाठकक लाभ स्पष्ट करत।

- ०-२० सेकेंड: पुस्तकक मुख्य प्रश्न।
- २०-६०: लेखक आ सन्दर्भ।
- १-२ मिनट: विषय आ संरचना।
- २-३: तीन विशेषता।
- ३-४: एक छोट उद्धरण वा दृश्यात्मक उदाहरण - उचित अधिकारसहित।
- ४-४: ४०: के पढ़थि आ किएक।
- ४: ४०-५: ००: स्रोत, पुस्तक-कड़ी आ समापन।

परियोजना-जाँच

- कथा वा पुस्तकक तथ्य शुद्ध अछि?
- कवर आ उद्धरणक उपयोग-अधिकार स्पष्ट अछि?
- जनरेटिव दृश्य पुस्तकक वास्तविक चित्रण जेकाँ भ्रामक नहि?
- स्वर स्वाभाविक मैथिलीमे अछि?
- पाँच मिनटमे अनावश्यक पुनरावृत्ति नहि?
- विवरणमे पुस्तक, लेखक, प्रकाशक आ स्रोत देल अछि?

सामान्य दोष आ समाधान

एहि अध्यायक बाद अहाँ -
- दृश्य, स्वर, अक्षर आ निर्यातक दोष चिन्हब
- लक्षित सुधार करब
- कखन पुन: जनाउ आ कखन सम्पादन करू, से तय करब

मुख्य शब्द: दोष-निदान · पुनर्जनन · मास्क · शोर · फ्लिकर

दोष	पहिल समाधान
चेहरा बदलि रहल	एके सन्दर्भ-चित्र, स्थिरता निर्देश, छोट क्लिप; असफल भाग काटू।
हाथ बिगड़ल	हाथ फ्रेमसँ बाहर, सरल क्रिया, नजदीकी दृश्य नहि; अन्य क्लिप चुनू।

वस्त्र बदलल	पात्र-पत्रक समान विवरण; image-to-video उपयोग।
कैमरा बहुत तेज	“locked camera” वा “very slow push-in”; अवधि बढ़ाउ।
दृश्य झिलमिल	कम गति, स्थिर पृष्ठभूमि, छोट क्लिप; सम्पादकमे स्थिरीकरण सीमित।
अक्षर गलत	जनरेटिव दृश्यसँ पाठ हटाउ; सम्पादकमे यूनिकोड पाठ जोड़ू।
मैथिली स्वर गलत	वाक्य छोट करू, ध्वन्यात्मक रूप, अपन रिकॉर्डिङ वा मानवीय सुधार।
सङ्गीत स्वर ढाकैत	सङ्गीत १०-१८ डेसीबल नीचाँ; स्वरक समय स्वतः ducking।
उपशीर्षक कटैत	सुरक्षित क्षेत्र, छोट फॉन्ट नहि, ९: १६ पूर्वावलोकन।
फाइल बहुत भारी	एच.२६४ एमपी४, उचित बिटरेट, अनावश्यक ४के नहि।

पुनः जनाउ कि सम्पादन करू?

पुनः जनाउ: मुख्य चेहरा, शरीर, वस्तु वा क्रिया मूलतः गलत।

सम्पादन करू: आरम्भ-अन्त खराब, अवधि लम्बा, रंग थोड़ा अलग, छोट पृष्ठभूमि दोष।

ढाकू/काटू: दोष एक क्षण वा फ्रेमक छोट भागमे।

त्यागू: क्लिपक अर्थ वा नैतिकता गलत; सुधारमे अधिक समय।

सुनहरा नियम दोषक नाम लिखने बिना फेर जनाउ नहि। “नीक नहि लागल” सुधारयोग्य निदान नहि अछि।

एकैस दिनक स्वयं-अध्ययन कार्यक्रम

एहि अध्यायक बाद अहाँ -

- नियमित अभ्याससँ कौशल बनायब
- तीन पूर्ण नमूना तैयार करब
- अपन उत्पादन-प्रणाली लिखब

मुख्य शब्द: दैनिक अभ्यास · नमूना-सङ्ग्रह · पुनरावलोकन · प्रकाशन

दिन	काज	पूरा
दिन १	ए.आइ. वीडियो प्रकार बुझू; पाँच उदाहरण चिन्हू।	□
दिन २	फोल्डर बनाउ; अनुपातक चार नमूना बनाउ।	□
दिन ३	एक मिनटक विचार आ दर्शक लिखू।	□
दिन ४	६०-९० सेकेंडक पटकथा।	□
दिन ५	छः दृश्यक दृश्य-पट्ट।	□
दिन ६	सात-अङ्गक पाँच प्रॉम्प्ट।	□
दिन ७	तीन पाठसँ वीडियो परीक्षण।	□
दिन ८	एक चित्रसँ तीन गति।	□
दिन ९	पात्र-पत्र आ निरन्तरता परीक्षण।	□
दिन १०	अपन स्वर रिकॉर्ड।	□
दिन ११	टीटीएसक छोट तुलना।	□

दिन १२	ध्वनि स्वच्छता आ सङ्गीत मिश्रण।	□
दिन १३	टाइमलाइन पर मोटा सम्पादन।	□
दिन १४	कट, संक्रमण आ रङ्ग-साम्य।	□
दिन १५	मैथिली उपशीर्षक।	□
दिन १६	३० सेकैंड परियोजना पूर्ण।	□
दिन १७	३ मिनट परियोजनाक पटकथा आ दृश्य-पट्ट।	□
दिन १८	दृश्य निर्माण आ सम्पादन।	□
दिन १९	प्रकाशन पैकेज, थम्बनेल आ विवरण।	□
दिन २०	कॉपीराइट, सहमति आ ए.आइ. प्रकटीकरण जाँच।	□
दिन २१	अन्तिम प्रकाशन, आत्म-मूल्याङ्कन आ अगिला लक्ष्य।	□

अन्तिम आत्म-मूल्याङ्कन

- □ विषय आ दर्शक स्पष्ट
- □ पटकथा समयमे
- □ दृश्य निरन्तर
- □ प्रॉम्प्ट सुरक्षित
- □ स्वर स्वाभाविक
- □ सङ्गीत संतुलित
- □ उपशीर्षक शुद्ध
- □ अधिकार/सहमति स्पष्ट
- □ ए.आइ. प्रकटीकरण उचित
- □ अन्तिम फाइलक बैकअप

अगिला लक्ष्य एकैस दिनक बाद कोनो एक विधा चुनू - पुस्तक-परिचय, शैक्षणिक पाठ, लोककथा, अभिलेखीय वृत्तचित्र वा सामाजिक लघु वीडियो - आ दस वीडियो धरि एके दृश्य-व्याकरण विकसित करू।

परिशिष्ट १ - तैयार प्रॉम्प्ट-सङ्ग्रह

नीचाँक प्रॉम्प्ट सभ आरम्भिक नमूना अछि। व्यक्ति, स्थान, समय, अनुपात आ शैली अपन परियोजना अनुसार बदलू। वास्तविक व्यक्ति वा स्थानक यथार्थ कृत्रिम दृश्य प्रकाशित करबाक पहिने सहमति आ प्रकटीकरण जाँचू।

1. मिथिला ग्रामक भोर

प्रॉम्प्ट A wide documentary shot of a Mithila village at dawn, mud courtyard, tulsi altar, soft mist, a woman drawing aripan, very slow camera push-in, natural colors, respectful realistic style, no text, no logo.

2. पोखरि आ मखान

प्रॉम्प्ट Aerial wide shot of a calm pond filled with makhana leaves, gentle morning breeze, small wooden boat

moving slowly, soft golden sunlight, realistic documentary style, stable camera.

3. मधुबनी चित्रमे सूक्ष्म गति

प्रॉम्प्ट Animate the original Madhubani painting while preserving flat colors, linework and composition; fish tails move gently, lotus leaves shimmer, no 3D transformation, slow camera push-in.

4. पुस्तक खुलब

प्रॉम्प्ट Close-up of an old Maithili book opening on a wooden desk, pages turning slowly in a soft breeze, warm window light, shallow depth of field, cinematic but realistic, no readable generated text.

5. लेखन दृश्य

प्रॉम्प्ट Top-down shot of a hand writing notes beside an open book, slow natural hand movement, warm study lamp, stable camera, documentary educational mood.

6. विद्यालय पाठ

प्रॉम्प्ट Bright classroom in rural Mithila, teacher pointing to a simple diagram, attentive students, medium shot, natural daylight, subtle camera movement, realistic educational documentary.

7. अभिलेखीय वातावरण

प्रॉम्प्ट Slow pan across carefully preserved old documents and family registers on a clean table, museum lighting, respectful archival documentary style, no invented readable text.

8. उत्पाद परिचय

प्रॉम्प्ट Clean studio shot of a handmade Mithila craft object rotating slowly on a neutral background, soft diffused light, accurate shape and color, commercial product video, no text.

9. वर्षा

प्रॉम्प्ट A locked-camera shot of monsoon rain falling in a Mithila courtyard, water drops from tiled roof, banana leaves moving gently, realistic ambience, soft gray light.

10. रेल यात्रा

प्रॉम्प्ट View through a train window moving across green fields in North Bihar, light rain on glass, calm travel documentary mood, stable realistic motion.

11. लोककथा वन

प्रॉम्प्ट Stylized 2D illustrated forest inspired by Mithila line art, moonlight, leaves moving softly, a deer pauses and listens, slow side tracking, clearly fictional animation.

12. सूर्योदय टाइमलैप्स

प्रॉम्प्ट Wide static shot of sunrise over fields, clouds and light changing gradually, smooth time-lapse feeling, natural colors, no sudden camera movement.

13. मोबाइल शॉर्ट हुक

प्रॉम्प्ट Vertical 9:16 macro close-up of makhana popping open during roasting, quick but clear action, dramatic natural sound, high detail, no text.

14. साक्षात्कार बी-रोल

प्रॉम्प्ट Medium close-up of an empty chair, microphone and books prepared for an interview, soft window light, very slow push-in, professional documentary mood.

15. प्रकाशन समापन

प्रॉम्प्ट A book resting beside a cup of tea and a lit lamp, evening ambience, slow pull-back revealing a quiet study, warm natural colors, space for title added later in editor.

16. स्थिर चित्रक गति

प्रॉम्प्ट Keep the person's face, clothing and background unchanged. Only the eyes blink naturally, the fabric moves slightly in the breeze, and the camera makes a very slow push-in.

17. पानीक प्रतिबिम्ब

प्रॉम्प्ट Close-up of water reflections with lotus leaves, gentle ripples spreading from a single drop, locked camera, soft morning light, meditative realistic mood.

18. हस्तकला हाथ

प्रॉम्प्ट Close-up documentary shot of artisan hands painting fine lines on a handmade object, accurate finger movement, stable camera, soft side light, no face shown.

19. मानचित्र विकल्प

प्रॉम्प्ट Minimal animated diagram showing a route with clean lines and simple icons on a plain background; no generated labels, leave empty spaces for text added in editing.

20. शीर्षक पृष्ठभूमि

प्रॉम्प्ट Subtle cream handmade paper texture with a thin burgundy Mithila-inspired border, very gentle light movement, empty center for title, no generated text.

परिशिष्ट २ - कार्यपत्रक आ जाँच-सूची

क. परियोजना-पत्र

विषय	लिखू
परियोजनाक नाम	
दर्शक	
मुख्य उद्देश्य	
अवधि	
मञ्च	
अनुपात	
मुख्य सन्देश	
आवश्यक स्रोत	
अधिकार/सहमति	
अन्तिम प्रकाशन दिनाङ्क	

ख. दृश्य-सूची

क्रम	समय	स्वर	दृश्य	प्रॉम्प्ट/स्रोत	स्थिति

ग. प्रकाशन पहिने जाँच

- □ शीर्षक सत्य आ स्पष्ट
- □ थम्बनेल वीडियोसँ मेल
- □ स्वर स्पष्ट
- □ सङ्गीत संतुलित
- □ मैथिली उपशीर्षक शुद्ध
- □ तथ्यक स्रोत
- □ चित्र/सङ्गीत लाइसेंस
- □ चेहरा/स्वर सहमति
- □ ए.आइ. प्रकटीकरण
- □ अन्तिम फाइल आ परियोजना बैकअप

परिशिष्ट ३ - वीडियो-शब्दावली

शब्द	अर्थ
अनुपात	चित्रक चौड़ाइ आ ऊँचाइक सम्बन्ध; उदाहरण १६: ९।
बी-रोल	मुख्य वाचन वा साक्षात्कारक ऊपर देखाओल सहायक दृश्य।
क्लिप	वीडियोक छोट अंश।
कट	एक दृश्यसँ दोसर दृश्य पर सीधा परिवर्तन।
डॉली	कैमरा भौतिक रूपसँ आगाँ वा पाछाँ चलब।
एफपीएस	प्रति सेकेंड फ्रेम संख्या; गति कतेक सुचारु देखायत।
फ्लिकर	प्रकाश, बनावट वा वस्तुक अनचाहा झिलमिल।
फ्रेम	वीडियोक एक स्थिर चित्र।
हुक	आरम्भक भाग जे दर्शकक ध्यान पकड़ैत अछि।
चित्रसँ वीडियो	स्थिर चित्रकेँ गति देनिहार जनरेटिव प्रक्रिया।
जे-कट	अगिला दृश्यक ध्वनि चित्र बदलनासँ पहिने शुरू।
एल-कट	पिछला दृश्यक ध्वनि अगिला चित्र पर जारी।
लिप-सिंक	ओठक गति आ स्वरक मेल।
मास्क	चित्रक चुनल भाग पर अलग सम्पादन।
पैन	कैमरा बाम-दहिन घूमब।
प्रॉम्प्ट	ए.आइ. औजारकेँ देल लिखित वा दृश्य निर्देश।
रेजोल्यूशन	चित्रमे पिक्सेलक मात्रा; उदाहरण १९२०×१०८०।
सन्दर्भ-चित्र	रूप, पात्र वा रचना बचाबय लेल देल चित्र।
सीड	किछु औजारमे पुनरावृत्त परिणाम प्रभावित करयवाला संख्या।
दृश्य-पट्ट	वीडियोक दृश्य-दर-दृश्य चित्रात्मक योजना।
एसआरटी	समय-सङ्केत सहित उपशीर्षकक पाठ-संचिका।
पाठसँ वीडियो	लिखित वर्णनसँ नूतन वीडियो क्लिप बनाबयक प्रक्रिया।
टाइमलाइन	सम्पादकमे समयक्रम अनुसार वीडियो, स्वर आ पाठ रखबाक क्षेत्र।

टिल्ट	कैमरा ऊपर-नीच घूमब।
संक्रमण	दू दृश्यक बीच परिवर्तनक प्रभाव।
टीटीएस	लिखित पाठ कें कृत्रिम स्वरमे बदलनाइ।
अपस्केल	वीडियो वा चित्रक रेजोल्यूशन बढ़ेनाइ।
वॉइस क्लोन	ककरो स्वर-सदृश कृत्रिम स्वर; स्पष्ट सहमति आवश्यक।

भाग २
विस्तारित व्याख्या आ अनुप्रयोग

ई पोथी ओहि सभ लेल अछि जे अपन बात वीडियोक माध्यमसँ कहऽ चाहैत छथि, मुदा जिनका लग ने कैमरा अछि, ने स्टूडियो, ने सम्पादनक प्रशिक्षण। कृत्रिम बुद्धि (ए.आइ.) आब ई तीनू काज एकटा साधारण मोबाइल वा कम्प्यूटरसँ सम्भव बना देलक अछि। जे बात कहियो लाखक-लाख टाकाक उपकरण आ दर्जनो लोकक टीमसँ होइत छल, से आब एकटा लिखल निर्देश — जकरा "प्रॉम्प्ट" कहल जाइत अछि — सँ भऽ सकैत अछि।

मुदा एकटा बात प्रारम्भहिमे स्पष्ट कऽ ली। ए.आइ. कोनो जादूक छड़ी नहि थिक। ओ एकटा औजार थिक — जेना कुरहरि, जेना कलम। कुरहरि काठ काटैत अछि, मुदा कोन गाछ काटब आ कोन घर बनाएब, से बढ़ई निर्णय करैत छथि। तहिना ए.आइ. दृश्य बना देत, स्वर बना देत, मुदा की कहबाक अछि, किएक कहबाक अछि, आ केकरा लेल कहबाक अछि — ई तीनू प्रश्नक उत्तर अहींक लग रहत। ई पोथी अहाँकेँ औजार चलाएब सिखाओत; कारीगरी अहाँक अपन रहत।

ई पोथी "स्वयं सीखू" शैलीमे लिखल गेल अछि — अर्थात् कोनो शिक्षक वा प्रशिक्षण-केन्द्रक आवश्यकता नहि। प्रत्येक अध्याय एकटा डेग थिक। पहिल अध्यायसँ प्रारम्भ करू, क्रमसँ आगाँ बढ़ू, आ प्रत्येक अध्यायक अन्तमे देल "अभ्यास" अवश्य करू। पढ़ब आ करब — दुनू संगे चलत तखने सीखब।

एकटा आर बात। एहि पोथीमे जतऽ चित्र देल गेल अछि, से सभ मौलिक निदर्शन-चित्र (इलस्ट्रेशन) थिक — कोनो कम्पनीक असली स्क्रीनक नकल नहि। कारण दूटा अछि: पहिल, ए.आइ. उपकरण सभक स्क्रीन प्रत्येक किछु मासमे बदलि जाइत अछि, तेँ असली स्क्रीनशॉट छपैत-छपैत पुरान भऽ जाइत; दोसर, सिद्धान्त बुझला पर कोनो स्क्रीन नव लागत, स्क्रीन रटला पर प्रत्येक बदलाव पर अहाँ अटकि जाएब। एहि पोथीक चित्र सभ ओहि सिद्धान्तकेँ देखबैत अछि जे सभ उपकरणमे समान रूपसँ लागू होइत अछि — प्रॉम्प्ट लिखबाक स्थान कतऽ रहैत अछि, सेटिङमे की-की रहैत अछि, टाइमलाइन केहन देखाइत अछि। एक बेर ई ढाँचा बुझि गेला पर अहाँ कोनो नव उपकरण अपने खोलि कऽ चला लेब — आ सैह "स्वयं सीखू" क असली अर्थ थिक।

मैथिलीमे तकनीकी पोथीक अभाव सदिखन खटकैत रहल अछि। हमर आग्रह अछि जे ई पोथी पढ़ि कऽ जे किछु सीखी, से मैथिलीहिमे वीडियो बना कऽ संसारकेँ देखाउ। मिथिलाक कथा, गीत, चित्रकला, इतिहास, पाबनि-तिहार — सभटा वीडियोक प्रतीक्षामे अछि।

ए.आइ. वीडियो की थिक?

कृत्रिम बुद्धि (आर्टिफिशियल इंटेलिजेंस, संक्षेपमे ए.आइ.) कम्प्यूटरक ओ क्षमता थिक जाहिसँ ओ मनुष्य-सन काज कऽ सकैत अछि — भाषा बुझब, चित्र चिन्हब, आ आब — चित्र तथा वीडियो बनाएब। जखन हम ए.आइ.केँ लिखि कऽ कही जे "पोखरिमे खेलाइत दूटा बच्चा देखाउ", आ ओ अपने ओहन दृश्यक चलचित्र बना दिअए, तँ ओकरा "जेनरेटिव ए.आइ." (सृजनशील ए.आइ.) कहल जाइत अछि। "जेनरेट" क अर्थ भेल — उत्पन्न करब। ई ए.आइ. कतहुसँ वीडियो ताकि कऽ नहि अनैत अछि; ओ नव वीडियो रचैत अछि, जे पहिने कतहु छलैके नहि।

ई केना सम्भव होइत अछि? सरल भाषामे बुझू। एहि ए.आइ. मॉडल सभकेँ करोड़ो चित्र आ वीडियो देखाओल गेल अछि। ओतबा देखला पर ओ सीखि गेल अछि जे बगुला केहन देखाइत अछि, पानि केना हिलैत अछि, भोरक इजोत केहन रंगक होइत अछि, आ चलैत मनुष्यक पएर केना उठैत अछि। जखन अहाँ लिखै छी "भोरमे पोखरि कात बगुला", तँ ओ अपन सीखल ज्ञानसँ ओहन दृश्य गढ़ि दैत अछि। ठीक जेना कोनो चित्रकार हजारो पोखरि देखने छथि आ आब बिना कोनो पोखरि सोझाँ रखने स्मृतिसँ पोखरि आँकि सकैत छथि।

ए.आइ.सँ वीडियो बनएबाक मुख्यतः चारि बाट अछि:

पहिल — पाठ-सँ-वीडियो (टेक्स्ट-टू-वीडियो) । अहाँ खाली लिखै छी, आ ए.आइ. पूरा दृश्य बना दैत अछि। ई सभसँ चमत्कारी बाट थिक, मुदा एहिमे नियन्त्रण किछु कम रहैत अछि — जे दृश्य आओत से ठीक-ठीक अहाँक कल्पना-जकाँ नहियो भऽ सकैत अछि।

दोसर — चित्र-सँ-वीडियो (इमेज-टू-वीडियो) । अहाँ एकटा स्थिर चित्र दै छी — फोटो वा ए.आइ.-निर्मित चित्र — आ ए.आइ. ओकरा चलायमान बना दैत अछि। एहिमे नियन्त्रण बेसी रहैत अछि, कारण प्रारम्भिक दृश्य अहाँ स्वयं चुनने छी।

तेसर — अवतार वीडियो। एकटा डिजिटल वक्ता (अवतार) स्क्रीन पर बैसल अहाँक लिखल पाठ बजैत अछि — जेना समाचार-वाचक। पाठ पढ़ाएब, घोषणा करब, वा व्याख्यान देबऽ लेल ई अत्यन्त उपयोगी अछि।

चारिम — ए.आइ.-सहायता-प्राप्त सम्पादन। एहिमे वीडियो अहाँक अपन रिकॉर्ड कएल रहैत अछि, मुदा काटब-जोड़ब, उपशीर्षक बनाएब, पृष्ठभूमि हटाएब — ई सभ काज ए.आइ. करैत अछि।

ई पोथी चारू बाट देखाओत, कारण व्यवहारमे नीक वीडियो प्रायः चारूक मिश्रणसँ बनैत अछि।

आब ई बुझू जे ए.आइ. की नहि कऽ सकैत अछि। ए.आइ. एखन धरि लम्बा वीडियो एक्के बेरमे नहि बनबैत अछि — प्रायः पाँचसँ बीस सेकेण्डक छोट-छोट क्लिप बनैत अछि, जकरा जोड़ि कऽ लम्बा वीडियो बनाओल जाइत अछि। ए.आइ.क बनाओल दृश्यमे कखनो काल भूल रहैत अछि — हाथक आङुर बेसी-कम भऽ जाइत अछि, लिखल अक्षर बिगड़ि जाइत अछि, एक क्लिपसँ दोसर क्लिपमे पात्रक मुँह बदलि जाइत अछि। आ सभसँ पैघ बात — ए.आइ.केँ मिथिला, मैथिली आ अहाँक हृदयक बातक कोनो ज्ञान नहि छैक; ओ ओतबे देत जतबा अहाँ माङब जनैत छी। तेँ एहि पोथीक आधा भाग "माङब केना" — अर्थात् योजना आ प्रॉम्प्ट — पर केन्द्रित अछि।

अभ्यास: कागज पर तीनटा एहन वीडियोक विषय लिखू जे अहाँ बनाबऽ चाहै छी। प्रत्येक लेल एक पंक्तिमे लिखू — ई वीडियो के देखत, आ देखला पर ओकरा की भेटत?

प्रारम्भिक तैयारी: की-की चाही?

ए.आइ.सँ वीडियो बनएबाक लेल कोनो महग यन्त्रक आवश्यकता नहि। जे चाही से एतबे —

पहिल, एकटा डिवाइस। स्मार्टफोन पर्याप्त अछि; कम्प्यूटर वा लैपटॉप रहत तँ आर सुविधा — पैघ स्क्रीन पर प्रॉम्प्ट लिखब, फाइल सम्हारब आ सम्पादन करब सहज होइत अछि। कोनो विशेष शक्तिशाली कम्प्यूटरक खगता नहि, कारण वीडियो अहाँक यन्त्र पर नहि, कम्पनीक सर्वर पर बनैत अछि। अहाँक यन्त्र खाली निर्देश पठबैत अछि आ बनल वीडियो डाउनलोड करैत अछि।

दोसर, इंटरनेट। जतेक नीक गति, ओतेक कम प्रतीक्षा। वीडियो फाइल पैघ होइत अछि, तेँ डाउनलोडमे डेटा खर्च होएत — मासमे किछु जी.बी.क अनुमान राखू।

तेसर, एकटा ईमेल खाता। प्रायः सभ ए.आइ. उपकरणमे खाता बनबऽ पड़ैत अछि, आ अधिकांशमे गूगल ईमेल (जीमेल) सँ सोझे प्रवेश भऽ जाइत अछि। एकटा सुझाव: अपन काजक लेल एकटा अलग ईमेल राखू, जाहिसँ ए.आइ. उपकरण सभक सूचना-पत्र अहाँक मुख्य ईमेलकेँ नहि भरए।

चारिम, "क्रेडिट"क समझ। अधिकांश ए.आइ. वीडियो उपकरण क्रेडिट-प्रणाली पर चलैत अछि। क्रेडिट एक प्रकारक कूपन थिक — एकटा वीडियो बनएबामे किछु क्रेडिट कटैत अछि। निःशुल्क खातामे थोड़ेक क्रेडिट भेटैत अछि — कतहु प्रतिदिन, कतहु मासमे, कतहु खाली एक्के बेर। सशुल्क योजनामे बेसी क्रेडिट, बेसी गुणवत्ता (जेना 1080पी वा 4के), बिना वॉटरमार्कक वीडियो, आ बेसी लम्बा क्लिप भेटैत अछि।

एतऽ एकटा व्यावहारिक नीति अपनाउ, जकरा हम "पहिने सस्त, फेर मस्त" कहै छी। कोनो नव उपकरण सिखबा काल निःशुल्क क्रेडिटसँ छोट-छोट प्रयोग करू। प्रॉम्प्ट लिखब सीखू, उपकरणक स्वभाव बुझू। जखन लागए जे आब गम्भीर काज करबाक अछि — जेना यूट्यूब चैनल चलाएब — तखने कोनो एकटा उपकरणक सशुल्क योजना लिअ। सभ उपकरणक योजना कीनब व्यर्थ; एकटा-दूटा पर्याप्त।

पाँचम, फाइल सम्हारबाक व्यवस्था। ई छोट बात लगैत अछि, मुदा अनुभवी लोक जनैत छथि जे बिना व्यवस्थाक काज कोना बिलाइत अछि। अपन कम्प्यूटर वा फोनमे एक-एक वीडियो परियोजनाक लेल अलग फोल्डर बनाउ। ओहिमे चारिटा उप-फोल्डर राखू — "पटकथा" (स्क्रिप्ट आ प्रॉम्प्ट), "क्लिप" (ए.आइ.सँ बनल कच्चा वीडियो), "स्वर-संगीत" (वॉयसओवर आ संगीत), आ "अन्तिम" (तैयार वीडियो) । प्रत्येक फाइलक नाम सार्थक राखू — "क्लिप-1-पोखरि-भोर" लिखल फाइल छह मास बादो चिन्हल जाएत, "वीडियो(3)-अन्तिम-नव-2" नहि।

आ अन्तमे — धैर्य। ए.आइ. उपकरण कखनो व्यस्त रहैत अछि, कखनो अजीब परिणाम दैत अछि, कखनो अहाँक नीक प्रॉम्प्टकेँ बुझबे नहि करैत अछि। ई सभ स्वाभाविक थिक। जे लोक तीन-चारि प्रयास धरि धैर्य रखैत छथि, से सीखि जाइत छथि; जे पहिले प्रयासक विचित्र परिणाम देखि छोड़ि दैत छथि, से वंचित रहि जाइत छथि।

अभ्यास: अपन डिवाइसमे उपरि देल फोल्डर-व्यवस्था बनाउ। एकटा नव ईमेल खाता (जँ चाही तँ) बनाउ। आ कोनो एकटा ए.आइ. वीडियो उपकरणक वेबसाइट खोलि कऽ खाली ई देखू जे निःशुल्क योजनामे की-की भेटैत अछि — एखन किछु बनाएब नहि अछि, खाली देखब अछि।

वीडियोक योजना: विचारसँ पटकथा धरि

चित्र १: ए.आइ. वीडियो निर्माणक छह चरण

सभसँ पैघ भूल जे नवसिखुआ करैत छथि से ई जे ओ सोझे उपकरण खोलि कऽ बटन दबाबऽ लगैत छथि। बिना योजनाक बनल वीडियो देखबामे ओहने लगैत अछि — बिना नक्शाक बनल घर। तेँ ई अध्याय उपकरणसँ पहिने अछि।

प्रत्येक वीडियोक पहिने तीन प्रश्नक उत्तर लिखू:

एक — उद्देश्य की? सिखाएब (जेना "मधुबनी चित्रकला केना बनैत अछि"), सुनाएब (जेना कोनो कथा), देखाएब (जेना गामक दृश्य), वा बजाएब (जेना कोनो घोषणा)? एक वीडियो, एक उद्देश्य — ई नियम मानू। जाहि वीडियोमे सभ किछु रहैत अछि, ताहिमे किछु नहि रहैत अछि।

दू — दर्शक के? बच्चा, विद्यार्थी, प्रवासी मैथिल, विदेशी जिज्ञासु, वा सामान्य दर्शक? दर्शके तय करत जे भाषा कतेक सरल हो, वीडियो कतेक लम्बा हो, आ दृश्य केहन हो। बच्चा लेल चमकैत रंग आ द्रुत गति; प्रौढ़ लेल ठहराओ आ गम्भीरता।

तीन — अवधि कतेक? प्रारम्भमे ३० सेकेण्डसँ २ मिनटक वीडियो बनाउ। छोट वीडियो बनाएब सहज, सुधारब सहज, आ आजुक दर्शककेँ पसिन्नो। मोन राखू — ए.आइ. पाँच-दस सेकेण्डक क्लिप बनबैत अछि, तेँ एक मिनटक वीडियोक अर्थ भेल छह-बारहटा क्लिप।

आब पटकथा (स्क्रिप्ट) । पटकथाक अर्थ कोनो जटिल साहित्यिक वस्तु नहि — ई खाली दू स्तम्भक तालिका थिक। बामा स्तम्भमे "की देखाएत" (दृश्य), दहिना स्तम्भमे "की सुनाएत" (स्वर वा उपशीर्षक) । तीस सेकेण्डक वीडियो लेल पाँच-छह पाँती पर्याप्त।

एतऽ ए.आइ. दोसर रूपमे सहायक बनि सकैत अछि — पटकथा-लेखक रूपमे। चैटबॉट (जेना क्लॉड, चैटजीपीटी, जेमिनी) केँ अपन विषय दियौ आ कहियौ जे पटकथा बना दिअए। मुदा माङबामे चतुराई चाही। खाली "मधुबनी पर वीडियो स्क्रिप्ट लिखू" कहला पर सामान्य, प्राणहीन पाठ भेटत। एना माङू:

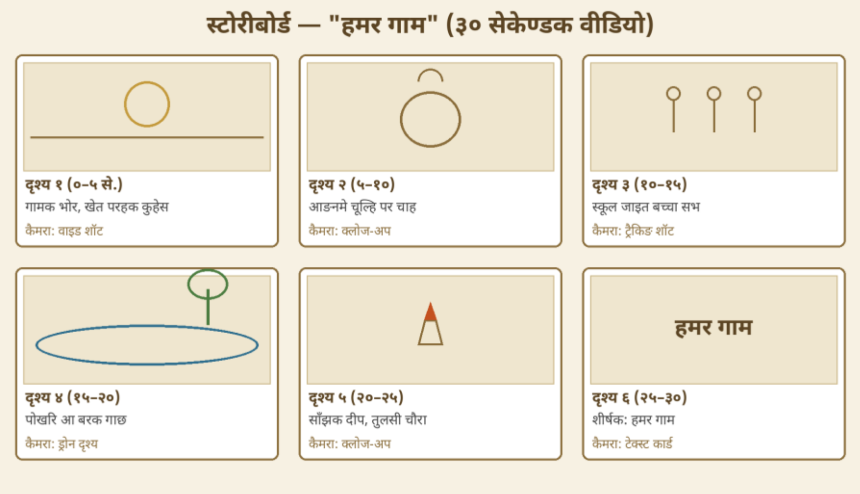
"हम मधुबनी चित्रकला पर ६० सेकेण्डक यूट्यूब वीडियो बना रहल छी। दर्शक: सामान्य भारतीय युवा जे मधुबनीक नाम सुनने अछि मुदा बेसी नहि जनैत अछि। स्वर: मैथिलीमे, सरल आ आत्मीय। हमरा दू स्तम्भमे पटकथा चाही — बामामे दृश्यक वर्णन (जकरा हम ए.आइ. वीडियो उपकरणकेँ देब), दहिनामे वॉयसओवरक पाठ। छह दृश्य, प्रत्येक दस सेकेण्डक।"

देखू — एहिमे विषय, अवधि, दर्शक, भाषा, स्वरूप आ उपयोग सभ किछु कहल गेल अछि। जतेक स्पष्ट माङ, ततेक नीक परिणाम। ई सूत्र आगाँ वीडियो-प्रॉम्प्टमे सेहो लागत।

चैटबॉटक बनाओल पटकथाकेँ आँखि मूनि कऽ स्वीकार नहि करू। ओकरा पढ़ू, अपन ज्ञानसँ जाँचू — तथ्य ठीक अछि? भाषा अहाँक अपन लगैत अछि? जे पाँती नीक नहि लागए से बदलू वा फेरसँ लिखबाउ ("तेसर दृश्य बेसी भावुक बनाउ", "अन्तिम पाँती छोट करू") । पटकथा अहाँक हस्ताक्षर थिक; ए.आइ. खाली मसिजीवी (लेखन-सहायक) थिक।

अभ्यास: अध्याय १ मे चुनल तीन विषयमेसँ एकटा लिअ। तीन प्रश्न (उद्देश्य, दर्शक, अवधि) क उत्तर लिखू। फेर कोनो चैटबॉटसँ उपरि देल ढंगक विस्तृत माङ कऽ कऽ दू-स्तम्भ पटकथा बनबाउ, आ ओहिमे कमसँ कम दूटा सुधार अपने हाथेँ करू।

स्टोरीबोर्ड: दृश्यक नक्शा



चित्र २: "हमर गाम" वीडियोक स्टोरीबोर्ड — प्रत्येक दृश्यक खाका, अवधि आ कैमरा-संकेत

पटकथा शब्दक नक्शा थिक; स्टोरीबोर्ड दृश्यक नक्शा। स्टोरीबोर्डमे वीडियोक प्रत्येक दृश्यक एकटा खाका — मोट-मोटी रेखाचित्र वा वर्णन — क्रमसँ राखल जाइत अछि। फिल्म-निर्माणमे ई सय बरखसँ चलि रहल विधि थिक, आ ए.आइ.-युगमे एकर महत्व घटल नहि, बढ़ल अछि। कारण? ए.आइ.केँ प्रत्येक क्लिप लेल अलग-अलग स्पष्ट निर्देश चाही, आ स्टोरीबोर्ड सैह निर्देश-सूची थिक।

चिन्ता नहि करू — चित्रकारी अएबाक कोनो आवश्यकता नहि। गोल मुँह, डण्टा-सन देह, आ तीर-निशान पर्याप्त। जँ रेखाचित्रो नहि बनाबऽ चाही तँ प्रत्येक दृश्यक तीन पाँतीक वर्णन लिखू — की देखाइत अछि, कैमरा की करैत अछि, आ कतेक सेकेण्ड।

कैमराक भाषा किछु सीखि लिअ, कारण यैह शब्द आगाँ प्रॉम्प्टमे काज देत:

वाइड शॉट — दूरसँ पूरा दृश्य; स्थान चिन्हएबा लेल। जेना गामक पूरा दृश्य।

क्लोज-अप — लग्गसँ; भाव आ बारीकी देखएबा लेल। जेना चूल्हि परहक चाहक भाप।

ट्रैकिङ शॉट — कैमरा पात्रक संग-संग चलैत अछि। जेना स्कूल जाइत बच्चाक पाछाँ।

ड्रोन दृश्य — ऊपरसँ, जेना चिड़ै देखैत अछि। जेना पोखरि आ खेतक विहंगम दृश्य।

डॉली इन/आउट — कैमरा धीरे-धीरे लग जाइत वा दूर होइत अछि; भावनात्मक जोर लेल।

स्थिर शॉट — कैमरा एक्के ठाम; शान्त, ठहरल दृश्य लेल।

दृश्य-क्रमक एकटा सरल सूत्र मोन राखू — "स्थापना, विस्तार, समापन"। पहिल दृश्य बताबए जे हम कतऽ छी (स्थापना — प्रायः वाइड शॉट)। बीचक दृश्य सभ विषयकेँ लग्गसँ देखाबए (विस्तार — क्लोज-अप, ट्रैकिङ)। अन्तिम दृश्य समेटए — कोनो भाव, कोनो सन्देश, वा शीर्षक-पट (समापन)।

चित्र २ मे "हमर गाम" नामक तीस सेकेण्डक वीडियोक स्टोरीबोर्ड देल अछि। देखू जे कोना छह दृश्यमे भोरसँ साँझ धरिक यात्रा अछि — भोरक कुहेस (स्थापना), चाह-स्कूल-पोखरि (विस्तार), दीप आ शीर्षक (समापन)। प्रत्येक दृश्यक संग अवधि आ कैमरा-संकेत लिखल अछि। यैह एक-एक दृश्य आगाँ चलि कऽ एक-एक प्रॉम्प्ट बनत।

स्टोरीबोर्ड बनएबा कालमे निरन्तरता (कन्टिन्युटी) पर ध्यान दिअ। जँ पहिल दृश्यमे भोर अछि तँ दोसरमे अचानक राति नहि हो। जँ कोनो पात्र लाल कुर्तामे अछि तँ सभ दृश्यमे लाल कुर्ता रहए — आ ई बात प्रत्येक प्रॉम्प्टमे लिखऽ पड़त, कारण ए.आइ.केँ पछिला क्लिप मोन नहि रहैत छैक। पात्रक वर्णन एकटा अलग कागज पर लिखि लिअ — "पात्र-पत्र" — आ प्रत्येक प्रॉम्प्टमे ओकरा जसक तस जोड़ू। यैह छोट चालि क्लिप-सभमे एकरूपता अनबाक सभसँ सरल उपाय थिक।

अभ्यास: अपन पटकथाकेँ छह दृश्यक स्टोरीबोर्डमे बदलू। प्रत्येक दृश्य लेल लिखू — दृश्य-वर्णन, कैमरा, अवधि। जँ कोनो पात्र अछि तँ ओकर "पात्र-पत्र" बनाउ (वस्त्र, अवस्था, रंग-रूप — तीन पाँतीमे)।

प्रॉम्प्ट लेखनक कला

चित्र ३: नीक प्रॉम्प्टक छह अंग — विषय, क्रिया, दृश्य, कैमरा, प्रकाश, शैली

प्रॉम्प्ट ओ लिखल निर्देश थिक जे अहाँ ए.आइ.केँ दै छी। ई ए.आइ.-युगक सभसँ मूल्यवान कौशल थिक, आ सुखद बात ई जे एहिमे कोनो तकनीकी ज्ञान नहि, भाषाक स्पष्टता चाही — आ भाषा तँ हमरा लोकनिक क्षेत्र थिक।

खराब प्रॉम्प्ट आ नीक प्रॉम्प्टक अन्तर देखू। खराब: "गामक दृश्य"। एहिसँ ए.आइ.केँ किछु नहि बुझाइत छैक — कोन देशक गाम? कोन ऋतु? दिन कि राति? की भऽ रहल अछि? ओ अपन मोनसँ किछु बना देत — प्रायः कोनो यूरोपीय वा अनाम गाम। आब नीक: "उत्तर भारतक एकटा गाम, भोरक समय, खेतक ऊपर हल्लुक कुहेस, फूसक आ खपड़ाक घर, दूरमे बरक गाछ, कैमरा धीरे-धीरे दहिना दिस घूमैत, नरम सुनहरा प्रकाश, यथार्थवादी वृत्तचित्र शैली।" आब ए.आइ. लग पूरा चित्र अछि।

चित्र ३ मे देखाओल छह अंगक सूत्र मोन राखू:

विषय — मुख्य के/की अछि? (दूटा बगुला)

क्रिया — ओ की कऽ रहल अछि? (माछ पकड़ैत)

दृश्य — कतऽ, कहिया? (कमलसँ भरल पोखरि, भोर)

कैमरा — कैमरा की करत? (धीरे आगाँ बढ़ैत)

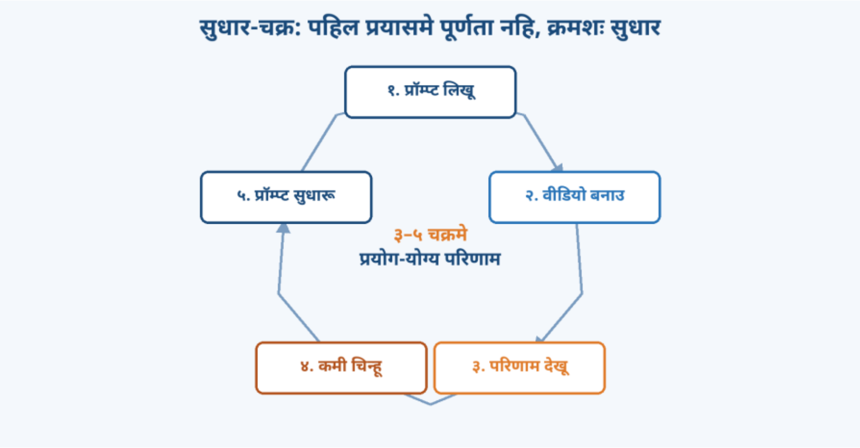
प्रकाश — केहन इजोत? (सुनहरा भोरक)

शैली — केहन रूप? (वृत्तचित्र, यथार्थवादी)

प्रत्येक प्रॉम्प्टमे ई छबो अंग रहए, त क्रम सेहो प्रायः यैह राखू — पहिने विषय-क्रिया, अन्तमे शैली। शैलीक किछु उपयोगी शब्द: यथार्थवादी (रियलिस्टिक), सिनेमाई (सिनेमैटिक), एनीमेशन, जलरंग चित्र-सन, पुरान फिल्म-सन, वृत्तचित्र (डॉक्युमेंट्री)।

भाषाक प्रश्न। अधिकांश ए.आइ. वीडियो उपकरण अंग्रेजी प्रॉम्प्ट सभसँ नीक बुझैत अछि, किछु हिन्दी सेहो बुझैत अछि। तँ व्यावहारिक बाट ई — पटकथा आ चिन्तन मैथिलीमे करू, आ अन्तिम प्रॉम्प्ट अंग्रेजीमे बनाउ। एहिमे चैटबॉट सहायक बनत: अपन मैथिली दृश्य-वर्णन चैटबॉटकेँ दिअ आ कहियौ — "एकर अंग्रेजी वीडियो-प्रॉम्प्ट बनाउ, छह अंग (विषय, क्रिया, दृश्य, कैमरा, प्रकाश, शैली) सहित।" एना अहाँक कल्पना मैथिलीक रहत, खाली अनुवाद मशीनी।

नकारात्मक प्रॉम्प्टक सेहो चलन अछि — जाहिमे लिखल जाइत अछि जे की नहि चाही। जेना: "धुँधला नहि, विकृत हाथ नहि, स्क्रीन पर कोनो लिखल पाठ नहि, वॉटरमार्क नहि।" किछु उपकरणमे एकर अलग स्थान रहैत अछि, किछुमे मुख्य प्रॉम्प्टहिमे जोड़ल जाइत अछि।



चित्र ४: सुधार-चक्र — प्रॉम्प्ट, परिणाम, समीक्षा, सुधार

आब सभसँ महत्वपूर्ण सिद्धान्त — सुधार-चक्र (चित्र ४) । पहिल प्रॉम्प्टसँ मनोनुकूल वीडियो बिरले भेटैत अछि, आ ई कोनो असफलता नहि थिक — यैह विधि थिक। परिणाम देखू, कमी चिन्हू, प्रॉम्प्टमे ओतबे बदलू, फेर बनाउ। बगुला बेसी पैघ आबि गेल? जोड़ू "छोट आकारक बगुला"। दृश्य बेसी चमकदम्म लागल? बदलू "नरम, हल्लुक प्रकाश"। प्रायः तीन-पाँच चक्रमे प्रयोग-योग्य क्लिप भेटि जाइत अछि। प्रत्येक चक्रमे एक-दुइए बात बदलू — सभटा एक्के बेर बदलला पर ई नहि बुझाएत जे कोन बदलाव काज कएलक।

अपन नीक प्रॉम्प्ट सभकेँ एकटा फाइलमे सञ्चित करैत चलू। छह मास बादक अहाँ आइक अहाँकेँ धन्यवाद देब।

अभ्यास: अपन स्टोरीबोर्डक पहिल दृश्य लेल छह-अंग सूत्रसँ एकटा पूर्ण प्रॉम्प्ट मैथिलीमे लिखू। फेर ओकर अंग्रेजी रूप (स्वयं वा चैटबॉटसँ) बनाउ। दुनूकेँ अपन "पटकथा" फोल्डरमे राखू।

पाठ-सँ-वीडियो उपकरण: पहिल क्लिप

चित्र ५: पाठ-सँ-वीडियो उपकरणक सामान्य ढाँचा (निदर्शन चित्र) — प्रॉम्प्ट-स्थान, सेटिङ, आ परिणाम

आब औजार हाथमे लेबाक क्षण आबि गेल। एहि अध्यायमे हम पाठ-सँ-वीडियो उपकरणक सामान्य ढाँचा बुझब आ पहिल क्लिप बनाएब।

पहिने उपकरण-परिचय। ई क्षेत्र अत्यन्त द्रुत गतिसँ बदलैत अछि — प्रत्येक किछु मासमे नव मॉडल अबैत अछि आ पुरनाक शक्ति बढ़ैत अछि। एहि पोथीक लिखबा काल (२०२६) प्रमुख नाम अछि: गूगलक वीओ (Veo), ओपनएआइक सोरा (Sora), क्लिङ (Kling), रनवे (Runway), पिक्सवर्स (PixVerse), सीडान्स (Seedance), हैलुओ (Hailuo), लूमा (Luma), पिका (Pika) । सभक स्वभाव किछु भिन्न — कियो यथार्थवादी दृश्यमे प्रबल, कियो शैलीगत/कलात्मक दृश्यमे, कियो गतिमे तेज आ दाममे सस्त। नाम सभ बदलैत रहत, मुदा ढाँचा — जे आगाँ लिखल अछि — सभमे लगभग समान रहैत अछि। तेँ नाम नहि, ढाँचा सीखू।

चित्र ५ देखू। लगभग प्रत्येक उपकरणमे यैह पाँच वस्तु भेटत:

एक — प्रॉम्प्ट-स्थान: एकटा पैघ खाली बॉक्स जतऽ अहाँ अपन निर्देश लिखब। ई उपकरणक हृदय थिक।

दू — मॉडल-चयन: एक्के कम्पनीक कतेको मॉडल रहैत अछि — नव आ शक्तिशाली (महग), पुरान वा "फास्ट" (सस्त) । सिखबा काल सस्त मॉडल चुनू; अन्तिम, प्रकाशन-योग्य क्लिप लेल नीक मॉडल।

तीन — अनुपात (आस्पेक्ट रेशियो): 16:9 यूट्यूब आ टीवी लेल (चौड़ा); 9:16 रील, शॉर्ट्स आ स्टेटस लेल (ठाढ़); 1:1 वर्गाकार। प्रारम्भहिमे सोचि लिअ जे वीडियो कतऽ जाएत, कारण बादमे अनुपात बदलला पर दृश्य कटैत अछि।

चारि — अवधि: प्रायः ५, ८, वा १० सेकेण्डक विकल्प। छोट क्लिप कम क्रेडिटमे बनत आ भूलो कम हएत।

पाँच — "बनाउ" बटन आ परिणाम-क्षेत्र: बटन दाबू, किछु सेकेण्डसँ किछु मिनट प्रतीक्षा करू, आ क्लिप परिणाम-क्षेत्रमे आबि जाएत। ओतहिसँ डाउनलोड करू।

आब पहिल क्लिपक विधि, डेग-दर-डेग:

१. उपकरणक वेबसाइट खोलू आ अपन ईमेलसँ खाता बनाउ।

२. निःशुल्क क्रेडिट जाँचू — कतेक अछि, कहिया नवीकृत होइत अछि।

३. मॉडलमे सस्त/फास्ट विकल्प, अनुपातमे 16:9, अवधिमे सभसँ छोट चुनू।

४. अध्याय ५ मे बनाओल अपन प्रॉम्प्ट साटू (पेस्ट करू) ।

५. "बनाउ" (Generate) दाबू आ प्रतीक्षा करू।

६. बनल क्लिपकेँ पूरा देखू — एक बेर नहि, दू-तीन बेर। हाथ, मुँह, अक्षर, चालि — सभ पर नजरि दिअ।

७. डाउनलोड कऽ अपन "क्लिप" फोल्डरमे सार्थक नामसँ राखू — भलहिं क्लिप अपूर्ण हो; तुलना लेल काज देत।

८. सुधार-चक्र चलाउ — कमी चिन्हू, प्रॉम्प्ट सुधारू, फेर बनाउ।

किछु व्यावहारिक टोटका। एक्के प्रॉम्प्टसँ दू-तीन बेर क्लिप बनाउ — ए.आइ. प्रत्येक बेर किछु भिन्न परिणाम दैत अछि, आ ताहिमेसँ नीक छाँटि लिअ। जाहि क्लिपमे मुख्य विषय ठीक अछि मुदा कात-करोटमे भूल अछि, से प्रायः सम्पादनमे काटल जा सकैत अछि — तुरत नहि फेकू। आ क्रेडिट सठबासँ पहिने दिनक प्रयोग-सीमा बना लिअ, नहि तँ उत्साहमे सप्ताहक क्रेडिट एक्के साँझ सठि जाएत।

अभ्यास: अपन स्टोरीबोर्डक पहिल दृश्यक क्लिप बनाउ — कमसँ कम तीन सुधार-चक्र चला कऽ। अन्तिम क्लिपक संग ओ प्रॉम्प्ट सेहो सुरक्षित करू जाहिसँ ओ बनल।

चित्रसँ वीडियो: बेसी नियन्त्रणक बाट

पाठ-सँ-वीडियोमे एकटा असुविधा अछि — प्रारम्भिक दृश्य पर अहाँक बस नहि चलैत अछि। ए.आइ. जे बनाओत, से बनाओत। एकर समाधान थिक चित्र-सँ-वीडियो (इमेज-टू-वीडियो): पहिने एकटा स्थिर चित्र तैयार करू जे ठीक अहाँक मोनक अछि, फेर ए.आइ.केँ कहू जे एहि चित्रकेँ चलायमान करू। एहिमे दृश्यक रूप-रंग, पात्रक चेहरा, वस्त्र — सभ पहिनहिसँ निश्चित रहैत अछि; ए.आइ. खाली गति जोड़ैत अछि।

चित्र कतऽसँ आओत? तीन स्रोत:

पहिल — अपन फोटो। गाम, पाबनि, प्रकृति, कला — अपन खिचल फोटोकेँ चलायमान करब सभसँ प्रामाणिक बाट थिक। मोन राखू, फोटो अहाँक अपन हो वा जकर फोटो थिक तकर अनुमति हो; आ जीवित व्यक्तिक फोटो चलायमान करबासँ पहिने हुनकर सहमति अवश्य लिअ — ई शिष्टाचारो थिक आ अध्याय १५ मे वर्णित नैतिकताक अंगो।

दोसर — ए.आइ.-निर्मित चित्र। चित्र बनएबाक अलग उपकरण अछि (आ प्रायः वीडियो-उपकरणहिमे चित्र-निर्माणक सुविधा रहैत अछि) । चित्रक प्रॉम्प्ट सेहो ओही छह-अंग सूत्रसँ बनत — खाली कैमराक गतिक स्थान पर चित्र-संयोजन (कम्पोजिशन) लिखू। चित्र सस्त आ द्रुत बनैत अछि, तेँ पहिने चित्रे पर सुधार-चक्र चलाउ — दस चित्र बना कऽ नीक छाँटब, दस वीडियो बनएबासँ बहुत सस्त पड़त।

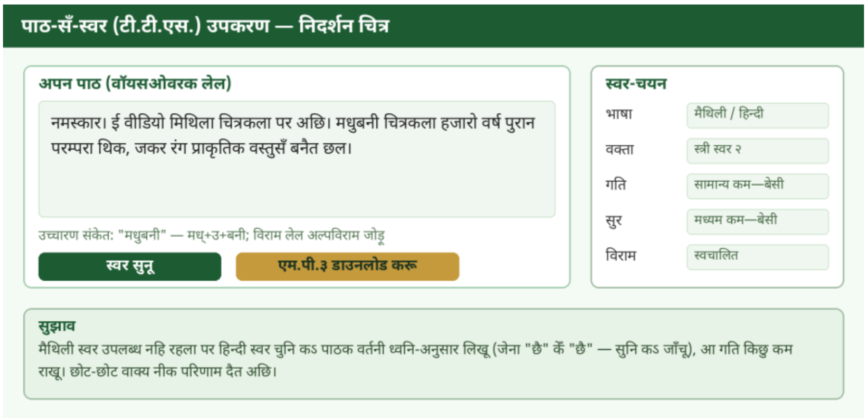
तेसर — अपन कलाकृति स्कैन कऽ कऽ। मिथिला चित्रकलाक कोनो कृतिकेँ (जँ अहाँक अपन बनाओल हो वा अधिकार हो) स्कैन कऽ हल्लुक गति देल जा सकैत अछि — जेना मछली हिलैत, रेखा-अलंकरण झिलमिलाइत। एहिमे अत्यन्त संयम राखू: परम्परागत कलाक गरिमा स्थिरतामे अछि, तेँ गति अत्यन्त सूक्ष्म राखू — "हल्लुक, धीमा, सूक्ष्म गति" (subtle motion) प्रॉम्प्टमे लिखू।

विधि सरल अछि। उपकरणमे "इमेज-टू-वीडियो" विकल्प चुनू, अपन चित्र अपलोड करू, आ संगमे गति-निर्देश लिखू — एहि प्रॉम्प्टमे आब दृश्य-वर्णन कम, गति-वर्णन बेसी रहत: की हिलत, कोन दिशामे, कतेक तेजीसँ, आ कैमरा की करत। जेना: "पानिमे हल्लुक लहरि, बगुलाक पाँखि धीरे-धीरे हिलैत, कैमरा अत्यन्त धीरे आगाँ बढ़ैत, आर किछु नहि बदलए।" अन्तिम वाक्यांश — "आर किछु नहि बदलए" — महत्वपूर्ण थिक; बिनु एकर ए.आइ. कखनो काल पूरा दृश्ये बदलि दैत अछि।

पात्र-एकरूपताक समस्याक ई सभसँ नीक समाधान थिक। जँ अहाँक वीडियोमे एक्के पात्र बेर-बेर आबए, तँ पहिने ओहि पात्रक एकटा उत्तम चित्र बनाउ (वा चुनू), आ प्रत्येक दृश्य लेल ओही चित्रकेँ भिन्न-भिन्न गति-निर्देशक संग चलायमान करू। किछु उन्नत उपकरणमे "क्यारेक्टर रेफरेन्स" वा "कन्सिस्टेन्ट क्यारेक्टर" नामक सुविधा रहैत अछि जाहिमे पात्रक चित्र एक बेर दऽ देला पर सभ क्लिपमे वैह पात्र अबैत अछि — अपन उपकरणमे ई सुविधा ताकू।

अभ्यास: कोनो चित्र (अपन फोटो वा ए.आइ.-निर्मित) लऽ कऽ ओकर तीन भिन्न-भिन्न गति-रूप बनाउ — एकमे खाली कैमरा चलए, दोसरमे खाली दृश्यक कोनो अंग हिलए, तेसरमे दुनू। तीनूक तुलना करू — कोन सभसँ स्वाभाविक लगैत अछि?

स्वर: वॉयसओवर आ पाठ-सँ-स्वर



चित्र ६: पाठ-सँ-स्वर (टी.टी.एस.) उपकरणक ढाँचा (निदर्शन चित्र)

वीडियोक आत्मा दृश्यमे नहि, स्वरमे बसैत अछि। दर्शक धुँधला दृश्य सहि लेत, मुदा खराब स्वर पर वीडियो बन्न कऽ देत। तेँ ई अध्याय ध्यानसँ पढ़ू — आ मैथिली-भाषी लेल एहिमे किछु विशेष बात अछि।

स्वर जोड़बाक दू बाट: अपन स्वर रिकॉर्ड करब, वा ए.आइ.सँ बनबाएब (पाठ-सँ-स्वर, अंग्रेजीमे टेक्स्ट-टू-स्पीच, संक्षेपमे टी.टी.एस.) ।

पहिने अपन स्वरक बात, कारण मैथिली लेल एखन धरि यैह सर्वोत्तम बाट थिक। मैथिलीक शुद्ध उच्चारण, स्वाभाविक लहजा, आ भावक उतार-चढ़ाव कोनो मशीन ओतेक नीक नहि आनि सकैत अछि जतेक मातृभाषी। आ रिकॉर्डिङ लेल स्टूडियोक खगता नहि — स्मार्टफोनक माइक आजुक दिनमे पर्याप्त नीक अछि। किछु नियम मानू: शान्त कोठलीमे रिकॉर्ड करू (पंखा बन्न, खिड़की बन्न, राति वा भोरक समय नीक); मुँहसँ फोन एक बीत्ता दूर राखू; ठाढ़ भऽ वा सोझ बैसि कऽ बाजू — स्वर खुजल रहत; पटकथा सोझाँ राखू मुदा पढ़ैत-सन नहि, कहैत-सन बाजू; आ एक्के बेरमे पूरा नहि, अनुच्छेद-अनुच्छेद रिकॉर्ड करू — भूल भेला पर खाली ओतबे फेर बाजू।

रिकॉर्ड कएल स्वरकेँ ए.आइ. निखारि सकैत अछि। कतेको उपकरण (सम्पादन-एपमे प्रायः "एन्हान्स वॉयस" वा "स्टूडियो साउण्ड" नामसँ) पृष्ठभूमिक शोर हटा कऽ स्वरकेँ स्टूडियो-सन बना दैत अछि। एकर प्रयोग अवश्य करू — साधारण रिकॉर्डिङ आ निखारल रिकॉर्डिङक अन्तर चकित करत।

आब पाठ-सँ-स्वर (चित्र ६) । एहिमे अहाँ पाठ लिखै छी, भाषा आ वक्ता (स्त्री/पुरुष, युवा/प्रौढ़) चुनै छी, गति-सुर मिलबै छी, आ एम.पी.३ फाइल डाउनलोड कऽ लै छी। हिन्दी, अंग्रेजी आदि लेल ई सुविधा अत्यन्त परिपक्व अछि। मैथिली लेल स्थिति सुधरि रहल अछि — किछु

उपकरण मैथिली स्वरक विकल्प देबऽ लागल अछि, आ भारतीय भाषाक लेल बनल उपकरण सभमे ई सम्भावना बेसी अछि। अपन उपकरणमे ताकू; जँ मैथिली भेटए तँ पहिने छोट पाठसँ जाँचू जे उच्चारण कतेक शुद्ध अछि।

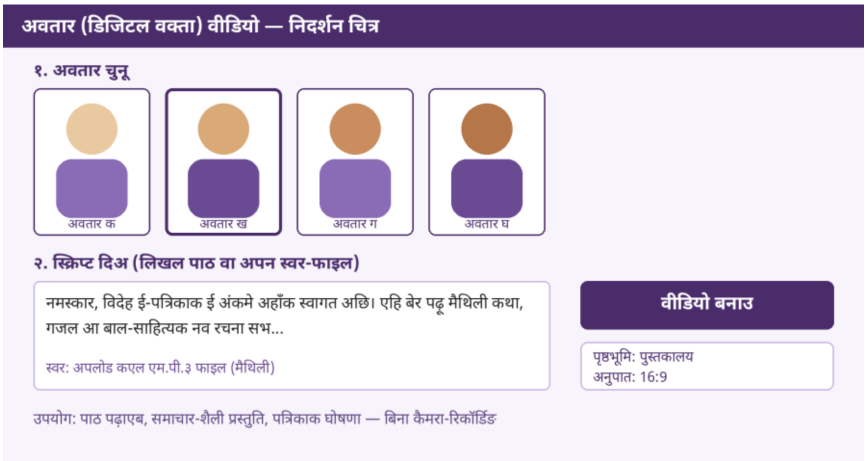
जँ मैथिली स्वर नहि भेटए तँ दूटा उपाय। पहिल आ उत्तम — अपन स्वर (उपरका विधिसँ) । दोसर — हिन्दी स्वरसँ काज चलाएब: पाठकेँ ध्वनि-अनुसार लिखू आ सुनि-सुनि कऽ वर्तनी मिलबैत जाउ; गति किछु धीमा राखू; छोट वाक्य बनाउ। परिणाम पूर्ण मैथिल लहजाक नहि हएत, मुदा काज-चलाउ भऽ जाएत। स्मरण राखू — ई समझौता थिक, आदर्श नहि; जतऽ भाव आ शुद्धता चाही (कविता, कथा, बाल-सामग्री) ततऽ अपने स्वर दिअ।

एकटा नव सुविधाक चर्चा — स्वर-अनुकरण (वॉयस क्लोनिङ) । किछु उपकरण अहाँक किछु मिनटक रिकॉर्डिङसँ अहाँक स्वरक डिजिटल प्रतिरूप बना दैत अछि, जे फेर कोनो पाठ अहाँक स्वरमे पढ़ि दैत अछि। मैथिली सामग्रीक लेल ई आकर्षक अछि — एक बेर अपन स्वर सिखा कऽ अनेक वीडियोक वॉयसओवर बनाओल जा सकैत अछि। मुदा दूटा कठोर नियम: केवल अपने स्वरक प्रतिरूप बनाउ — आन केकरो स्वरक अनुकरण बिना लिखित अनुमति सर्वथा वर्जित; आ जाहि वीडियोमे प्रतिरूप-स्वरक प्रयोग हो ताहिमे एकर सूचना देब नीक अभ्यास थिक।

स्वर आ दृश्यक मेल अन्तमे सम्पादनमे हएत (अध्याय ११), तेँ स्वर-फाइल अलगसँ "स्वर-संगीत" फोल्डरमे राखू। वॉयसओवर पहिने बनाउ, वीडियो-क्लिप पाछाँ — ई क्रम नीक, कारण स्वरक लम्बाइ सुनि कऽ बुझाइत अछि जे कोन दृश्य कतेक सेकेण्डक चाही।

अभ्यास: अपन पटकथाक वॉयसओवर दुनू विधिसँ बनाउ — एक बेर अपन स्वरमे रिकॉर्ड कऽ, एक बेर कोनो टी.टी.एस.सँ। दुनूकेँ आँखि मूनि कऽ सुनू। कोन सुनबामे बेसी "अपन" लगैत अछि? किएक?

अवतार वीडियो: डिजिटल वक्ता



चित्र ७: अवतार वीडियो उपकरणक ढाँचा (निदर्शन चित्र) — अवतार-चयन आ स्क्रिप्ट

कल्पना करू — अहाँकेँ प्रत्येक पन्द्रह दिन पर पत्रिकाक नव अंकक घोषणा-वीडियो बनएबाक अछि, वा कोनो पाठ्यक्रमक पचास व्याख्यान। प्रत्येक बेर कैमरा, प्रकाश, वस्त्र, रिकॉर्डिङ — ई सम्भव नहि। एकर समाधान थिक अवतार वीडियो: एकटा डिजिटल मनुष्य स्क्रीन पर बैसल अहाँक लिखल पाठ बजैत अछि — ठोर हिलैत, आँखि झपकैत, हाथ चलैत — जेना समाचार-वाचक।

एहि श्रेणीक प्रसिद्ध उपकरण अछि हेजेन (HeyGen), सिन्थेसिया (Synthesia) आदि — आ ई श्रेणी सेहो द्रुत गतिसँ बढ़ि रहल अछि। ढाँचा सभमे लगभग समान (चित्र ७):

एक — अवतार चुनू। उपकरणमे सैकड़ो बनल-बनाओल अवतार रहैत अछि — भिन्न वयस, वेश, मुद्राक। किछु उपकरणमे अपन फोटो वा छोट वीडियोसँ अपने अवतार बनएबाक सुविधा सेहो अछि — अर्थात् स्क्रीन पर "अहीं" रहब, बिना प्रत्येक बेर रिकॉर्ड कएने। (अपने अवतार बनाबी तँ ओही सहमति-नियमक पालन करू जे स्वर-अनुकरणमे कहल गेल — केवल अपन छवि, आनक नहि।)

दू — स्क्रिप्ट दिअ। दू रूपमे भऽ सकैत अछि: लिखल पाठ (जकरा उपकरण टी.टी.एस.सँ बजाओत) वा अपन रिकॉर्ड कएल स्वर-फाइल (जकरा अवतार ठोर मिला कऽ बाजत) । मैथिली लेल दोसर बाट प्रायः नीक — अपन मैथिली वॉयसओवर अपलोड करू, अवतार ओकरे बाजत। ठोरक मेल (लिप-सिंक) आश्चर्यजनक रूपेँ नीक भेटत।

तीन — पृष्ठभूमि आ ढाँचा चुनू — पुस्तकालय, कार्यालय, सादा रंग, वा अपन कोनो चित्र। अनुपात चुनू (16:9 वा 9:16), आ बनाउ।

अवतार वीडियोक शोभा एकर व्यावहारिकतामे अछि, चमत्कारमे नहि। समाचार-शैली प्रस्तुति, पत्रिका-घोषणा, पाठ-व्याख्या, संस्था-परिचय, सूचना-वीडियो — एहि सभमे ई अत्युत्तम। कथा-कविताक भाव-प्रधान प्रस्तुतिमे एखनो मनुष्ये नीक।

एकटा सुझाव प्रस्तुतिक: अवतारकेँ अनवरत पूरा वीडियो भरि नहि राखू। बीच-बीचमे सम्बन्धित दृश्य, चित्र, वा पाठ-पट (जकरा "बी-रोल" कहल जाइत अछि) देखाउ आ अवतारक स्वर पाछाँ चलैत रहए। एहिसँ वीडियो जीवन्त लागत। ई जोड़-तोड़ सम्पादनमे हएत — अवतार-क्लिप आ बी-रोल क्लिप दुनू अलग-अलग बना कऽ राखू।

आ हँ — जखन वीडियोमे अवतार मनुष्य-सन देखाइत हो, तखन दर्शककेँ ई बुझबाक अधिकार छैक जे ई डिजिटल वक्ता थिक। विवरणमे एक पाँती लिखि दिअ। एहिसँ विश्वास बनल रहत — आ विश्वासे चैनलक असली पूँजी थिक।

अभ्यास: कोनो अवतार-उपकरणक निःशुल्क योजनामे ३०-सेकेण्डक एकटा परिचय-वीडियो बनाउ — विषय: "हमर गामक परिचय" वा अपन कोनो प्रिय पोथीक परिचय। लिखल पाठ आ अपलोड कएल स्वर — दुनू विधि आजमाउ।

संगीत आ ध्वनि-प्रभाव

दृश्य आँखि लेल, स्वर कान लेल, आ संगीत — हृदय लेल। ओहने दृश्य बिना संगीतक सुन्न लगैत अछि आ उचित संगीतक संग जीबि उठैत अछि। मुदा संगीतक संग सभसँ पैघ खतरा अछि — कॉपीराइट। ककरो गीत बिना अनुमतिक वीडियोमे लगओला पर यूट्यूब वीडियोकेँ रोकि सकैत अछि, आय आन कियो लऽ सकैत अछि, आ चैनल पर दण्ड लागि सकैत अछि। तेँ नियम एक: प्रसिद्ध फिल्मी वा रिकॉर्ड कएल गीत कहियो नहि — जा धरि लिखित अनुमति नहि।

तखन संगीत आओत कतऽसँ? तीन वैध स्रोत:

पहिल — ए.आइ.-निर्मित संगीत। आब एहन उपकरण अछि जे लिखल वर्णनसँ संगीत बना दैत अछि — "धीमा, भावुक, बाँसुरी-प्रधान, ६० सेकेण्ड" लिखू आ संगीत तैयार। किछु उपकरण गीतक बोलसँ पूरा गीत (स्वर सहित) सेहो बनबैत अछि। अपन वीडियोक लेल बनाओल एहन संगीत प्रयोगमे प्रायः सुरक्षित अछि, मुदा प्रत्येक उपकरणक उपयोग-शर्त (लाइसेन्स) एक बेर पढ़ि लिअ — विशेषतः ई जे व्यावसायिक प्रयोग (यूट्यूब मुद्रीकरण सहित) क अनुमति अछि कि नहि।

दोसर — कॉपीराइट-मुक्त संगीत-पुस्तकालय। यूट्यूबक अपन "ऑडियो लाइब्रेरी" (यूट्यूब स्टूडियोक भीतर) निःशुल्क आ सुरक्षित अछि। एकर अतिरिक्त कतेको वेबसाइट मुक्त-लाइसेन्स संगीत दैत अछि — किछु पूर्ण निःशुल्क, किछु "श्रेय दिअ" (एट्रिब्यूशन) शर्त पर। श्रेय-शर्त होअए तँ वीडियोक विवरणमे संगीतकारक नाम लिखब नहि बिसरू।

तेसर — अपन रिकॉर्ड कएल संगीत। मिथिलाक अपन समृद्ध संगीत-परम्परा अछि। जँ अहाँ वा अहाँक परिचित कियो गबैत-बजबैत छथि, तँ अपने रिकॉर्ड कएल लोकधुन सभसँ प्रामाणिक स्रोत थिक — आ अहाँक वीडियोकेँ ओ पहिचान देत जे कोनो ए.आइ. नहि दऽ सकैत अछि। (लोकगीतक धुन परम्परागत अछि, मुदा ककरो विशेष रिकॉर्डिङ वा रचना हुनक अपन — ई भेद मोन राखू।)

संगीत लगएबाक कला: वॉयसओवरक पाछाँ संगीत मद्धिम राखू — मुख्य स्वरक बीस-तीस प्रतिशत। जतऽ वॉयसओवर नहि (प्रारम्भ, अन्त, दृश्य-परिवर्तन) ततऽ संगीत उठा सकै छी। वीडियोक भावक संग संगीतक भाव मिलए — भोरक दृश्यमे उल्लास, विदाइक दृश्यमे करुणा। आ अन्तमे संगीत धीरे-धीरे मद्धिम भऽ कऽ बिला जाए (फेड आउट) — झटसँ बन्न भेल संगीत कानकेँ धक्का दैत अछि।

ध्वनि-प्रभाव (साउण्ड इफेक्ट) छोट-छोट ध्वनि थिक — चिड़ैक चहचह, पानिक छपाक, हवाक सरसराहट। ई दृश्यकेँ विश्वसनीय बनबैत अछि। किछु नव वीडियो-मॉडल दृश्यक संग ध्वनि सेहो बना दैत अछि ("नेटिव ऑडियो") — अपन उपकरणमे ई सुविधा हो तँ सक्रिय राखू; नहि तँ मुक्त ध्वनि-पुस्तकालयसँ लऽ कऽ सम्पादनमे जोड़ू।

अभ्यास: अपन वीडियो लेल दू भिन्न शैलीक पृष्ठ-संगीत जुटाउ — एकटा ए.आइ.-निर्मित, एकटा मुक्त-पुस्तकालयसँ। दुनूकेँ वीडियोक पाछाँ (मनहिमे) बजा कऽ सोचू — कोन भाव बेसी मिलैत अछि?

सम्पादन: क्लिप सभकेँ वीडियो बनाएब

चित्र ८: सम्पादकक टाइमलाइन (निदर्शन चित्र) — वीडियो, स्वर, संगीत आ उपशीर्षकक पट्टी

आब अहाँ लग सभ सामग्री अछि — वीडियो-क्लिप, वॉयसओवर, संगीत। सम्पादन ओ रसोई थिक जतऽ ई सभ सामग्री मिलि कऽ व्यंजन बनत। आ नीक बात ई जे सम्पादन-कौशल एक बेर सीखि लेला पर सभ वीडियोमे काज देत — ई ए.आइ.-उपकरण-सन बदलैत नहि रहैत अछि।

सम्पादक (एडिटर) चुनू। मोबाइल आ कम्प्यूटर दुनू लेल नि:शुल्क सम्पादक उपलब्ध अछि — कैपकट (CapCut) एहि समय सभसँ प्रचलित आ सरल अछि; कम्प्यूटर पर डा-विन्ची रिजॉल्व (DaVinci Resolve) नि:शुल्क रहितो व्यावसायिक स्तरक अछि। कोनो चुनू — ढाँचा (चित्र ८) सभमे समान:

मीडिया-क्षेत्र: जतऽ अहाँ अपन सभ फाइल आनब (इम्पोर्ट) । पूर्वावलोकन: जतऽ वीडियो चलि कऽ देखाइत अछि। टाइमलाइन: सभसँ महत्वपूर्ण — समयक रेखा, जाहि पर क्लिप सभ क्रमसँ सजाओल जाइत अछि। टाइमलाइनमे कतेको पट्टी (ट्रैक) रहैत अछि — एकटामे वीडियो, एकटामे स्वर, एकटामे संगीत, एकटामे उपशीर्षक — सभ एक-दोसराक ऊपर, संगे-संग बजैत।

सम्पादनक मूल क्रम एना राखू:

१. पहिने वॉयसओवर टाइमलाइन पर राखू। यैह वीडियोक रीढ़ थिक — दृश्य एकरे अनुसार सजत।

२. स्वर सुनैत-सुनैत वीडियो-क्लिप सभ क्रमसँ राखू — जे बात बाजल जा रहल अछि, वैह दृश्य देखाए। क्लिपकेँ काटू (कट) — प्रत्येक क्लिपक सभसँ नीक अंश राखू, शेष हटाउ। ए.आइ.-क्लिपक प्रारम्भ आ अन्तमे प्राय: अटपटाहटि रहैत अछि — दुनू छोर छाँटि देला पर क्लिप निखरि जाइत अछि।

३. संक्रमण (ट्रांजिशन) जोड़ू — एक क्लिपसँ दोसरमे जएबाक ढंग। नियम: जतेक कम, ततेक नीक। सोझ कट सभसँ शुद्ध; भाव-परिवर्तन पर हल्लुक फेड। घुमैत-नचैत रंगीन संक्रमण नवसिखुआक चिन्ह थिक।

४. संगीत-पट्टीमे संगीत राखू, स्वर मद्धिम करू (अध्याय १०) ।

५. रंग-सुधार: अधिकांश सम्पादकमे एक-क्लिक "फिल्टर" वा "एन्हान्स" अछि। ए.आइ.-क्लिप सभ भिन्न-भिन्न उपकरणसँ आएल हो तँ सभ पर एक्के फिल्टर लगाउ — रंगक एकरूपता आबि जाएत, आ वीडियो सिलल-सन लागत।

६. प्रारम्भमे शीर्षक-पट (३-४ सेकेण्ड) आ अन्तमे समापन-पट (चैनलक नाम, "धन्यवाद") जोड़ू।

निर्यात (एक्सपोर्ट) क सेटिङ: 1080पी, एम.पी.४ प्रारूप, ३० फ्रेम प्रति सेकेण्ड — ई यूट्यूब लेल मानक अछि। निर्यात कएल फाइल "अन्तिम" फोल्डरमे राखू।

एकटा सुझाव — पूरा वीडियो बना कऽ एक बेर ओकरा आद्योपान्त बिना रोकने देखू, दर्शक बनि कऽ। जतऽ-जतऽ मोन उचटए, बुझू ओतऽ काट-छाँट चाही। फेर कोनो घरक सदस्यकेँ देखाउ — पहिल दर्शकक प्रतिक्रिया सय समीक्षकसँ बेसी सिखबैत अछि।

अभ्यास: अपन सभ क्लिप, स्वर आ संगीतकेँ जोड़ि कऽ पहिल पूर्ण वीडियो बनाउ — शीर्षक-पट आ समापन-पट सहित। निर्यात कऽ कऽ परिवारक कोनो सदस्यकेँ देखाउ आ हुनकर पहिल प्रतिक्रिया लिखि लिअ।

उपशीर्षक: पढ़बा-योग्य वीडियो

चित्र ९: एस.आर.टी. उपशीर्षक फाइलक स्वरूप

आजुक अधिकांश दर्शक वीडियो बिना स्वरक देखैत छथि — बसमे, कार्यालयमे, राति बिछौन पर। उपशीर्षक (सबटाइटल) नहि तँ दर्शक नहि। आ मैथिली सामग्री लेल उपशीर्षकक महत्व दोबर अछि: मैथिली उपशीर्षकसँ भाषा पढ़बाक अभ्यास बढ़ैत अछि, आ हिन्दी/अंग्रेजी उपशीर्षकसँ ओहो दर्शक जुड़ैत छथि जे मैथिली नहि बुझैत छथि — अर्थात् अहाँक मिथिलाक कथा संसार भरि पहुँचैत अछि।

उपशीर्षकक मानक प्रारूप थिक एस.आर.टी. (SRT) — एकटा साधारण पाठ-फाइल जाहिमे तीन वस्तु बेर-बेर अबैत अछि (चित्र ९): क्रम-संख्या, समय-रेखा (कोन सेकेण्डसँ कोन सेकेण्ड धरि), आ पाठ। ई फाइल कोनो साधारण पाठ-सम्पादक (नोटपैड) मे सेहो बनल-सुधारल जा सकैत अछि।

मुदा हाथसँ समय मिलाएब श्रमसाध्य अछि, आ एतऽ ए.आइ. फेर सहायक। दू बाट:

पहिल — स्वचालित प्रतिलेखन (ऑटो ट्रान्सक्रिप्शन) । सम्पादक (जेना कैपकट) मे "ऑटो कैप्शन" सुविधा वीडियोक स्वर सुनि कऽ अपने उपशीर्षक बना दैत अछि, समय सहित। हिन्दी-अंग्रेजीमे ई अत्यन्त शुद्ध अछि; मैथिली स्वरकेँ प्रायः हिन्दी बुझि कऽ लिखत — तखन बनल पाठकेँ अहाँ सुधारब। सुधारबो हाथसँ लिखबासँ बहुत तेज होइत अछि, कारण समय-रेखा बनल-बनाओल भेटैत अछि।

दोसर — पटकथासँ। अहाँ लग तँ पटकथा पहिनहिसँ लिखल अछि! चैटबॉटकेँ अपन पटकथा आ वीडियोक कुल अवधि दऽ कऽ कहू — "एकरा एस.आर.टी. प्रारूपमे बाँटू, प्रत्येक उपशीर्षक अधिकतम दू पंक्ति, पढ़बाक सहज गतिसँ।" बनल फाइलक समय-रेखा सम्पादकमे राखि कऽ आगाँ-पाछाँ सरकाउ।

उपशीर्षकक शिल्प-नियम: एक बेरमे अधिकतम दू पंक्ति, प्रति पंक्ति लगभग बत्तीस-चालीस अक्षर। पाठ स्क्रीन पर कमसँ कम एक सेकेण्ड, वाक्य-खण्ड अर्थपूर्ण ठाम टूटए ("हम गेलौं / बजार" नहि, "हम बजार गेलौं" एक्के संग) । अक्षरक रंग उज्जर, पाछाँ हल्लुक कारी छाया वा पट्टी — जाहिसँ चमकैत दृश्यो पर पढ़ल जा सकए। स्क्रीनक निच्चाँ मध्यमे राखू, मुदा एतेक निच्चाँ नहि जे रील-प्रारूपमे कटि जाए।

यूट्यूब पर उपशीर्षक दू तरहेँ देल जा सकैत अछि — वीडियोमे सटल ("बर्न्ड-इन", जे सम्पादकसँ निर्यातमे जुड़ि जाइत अछि) वा अलग एस.आर.टी. फाइल रूपमे अपलोड कएल (जकरा दर्शक चालू-बन्न कऽ सकैत छथि) । उत्तम विधि: मैथिली उपशीर्षक अलग फाइलसँ दिअ, आ हिन्दी/अंग्रेजी सेहो अलग-अलग एस.आर.टी. बना कऽ जोड़ू — अनुवादमे चैटबॉट सहायता करत (अनुवाद जँचबाक दायित्व अहाँक) । एना एक्के वीडियो तीन भाषाक दर्शक धरि पहुँचत।

अभ्यास: अपन वीडियोक मैथिली एस.आर.टी. बनाउ (कोनो एक विधिसँ), फेर ओकर हिन्दी अनुवाद-फाइल बनाउ। दुनूकेँ वीडियोक संग चला कऽ जाँचू — समय ठीक अछि? पंक्ति कतहु बेसी लम्बा तँ नहि?

यूट्यूब पर प्रकाशन

चित्र १०: यूट्यूब अपलोड जाँच-सूची

वीडियो बनि गेल; आब ओकरा दर्शक धरि पहुँचाएब। यूट्यूब एखनो सभसँ पैघ आ स्थायी मंच थिक — एतऽ राखल वीडियो बरख-दर-बरख देखल जाइत रहैत अछि, जखन कि रील-प्रारूपक मंच पर वीडियोक आयु किछुए दिन। तेँ मुख्य घर यूट्यूबकेँ बनाउ; रील-शॉर्ट्स ओकर खिड़की।

चैनल बनाएब सरल — यूट्यूब पर अपन गूगल खातासँ प्रवेश करू आ चैनल बनाउ। चैनलक नाम सोचि कऽ राखू: छोट, उच्चारण-सहज, आ विषय-सूचक। चैनल-चित्र (लोगो) आ बैनर सेहो ए.आइ. चित्र-उपकरणसँ बना सकै छी।

अपलोडक समय चित्र १० क जाँच-सूची सोझाँ राखू। किछु बिन्दुक विस्तार:

शीर्षक: पहिल तीन-चारि शब्दमे मुख्य बात। मैथिली शीर्षकक संग कोष्ठकमे हिन्दी वा अंग्रेजी जोड़ला पर खोजमे वीडियो बेसी भेटैत अछि — कारण खोजनिहार सभ भाषामे खोजैत छथि।

विवरण: पहिल दू पंक्ति सभसँ मूल्यवान — यैह खोज-परिणाममे देखाइत अछि। एहिमे वीडियोक सार लिखू, मुख्य शब्द सहित। नीचाँ विषय-क्रम (कोन मिनट पर की अछि), स्रोत-सूची, आ चैनल-परिचय।

थम्बनेल (मुख-चित्र): दर्शक शीर्षकसँ पहिने थम्बनेल देखैत छथि। नियम: एक चित्र, एक भाव, अधिकतम तीन शब्द — आ शब्द एतेक पैघ जे मोबाइलक छोट स्क्रीन पर पढ़ाए। ए.आइ. चित्र-उपकरणसँ आकर्षक थम्बनेल बनाउ, मुदा भ्रामक नहि — जे थम्बनेलमे अछि से वीडियोमे होएबाक चाही, नहि तँ दर्शक ठकल अनुभव करत आ चैनलसँ विश्वास उठि जाएत।

ए.आइ.-घोषणा: यूट्यूब आब अपेक्षा करैत अछि जे यथार्थ-सन देखाइत ए.आइ.-निर्मित वा परिवर्तित सामग्रीक सूचना अपलोडक समय देल जाए ("Altered content" प्रश्नक उत्तरमे) । ई खाली नियम-पालन नहि, दर्शकक संग ईमानदारी थिक। एनिमेशन-सन स्पष्टतः कल्पित शैलीमे ई बाध्यता प्रायः नहि, मुदा शंका हो तँ घोषित कऽ देब सदैव नीक।

भाषा-सेटिङ: वीडियोक भाषा "मैथिली" चुनू (यूट्यूबक सूचीमे अछि) । एहिसँ मैथिली खोजनिहार धरि वीडियो पहुँचबामे सहायता होइत अछि, आ मैथिली सामग्रीक आँकड़ो सुदृढ़ होइत अछि।

अपलोडक बाद की? पहिल घण्टा-दिनक प्रतिक्रिया देखू, टिप्पणीक उत्तर दिअ — प्रारम्भिक संवाद चैनलक जड़ि सींचैत अछि। आ नियमितता राखू: पाक्षिक एक वीडियो सालक चौबीस वीडियो — ई अनियमित सय वीडियोसँ बेसी फल दैत अछि। दर्शक कार्यक्रमसँ जुड़ैत छथि, छिटपुटसँ नहि।

रील/शॉर्ट्स लेल अपन मुख्य वीडियोक सभसँ रोचक तीस-साठि सेकेण्ड 9:16 अनुपातमे काटू (सम्पादकमे "रीफ्रेम" वा क्रॉप सुविधासँ), अन्तमे लिखू — "पूरा वीडियो चैनल पर।" यैह खिड़की नव दर्शककेँ घर धरि आनत।

अभ्यास: अपन वीडियो अपलोड करू — जाँच-सूचीक सातू बिन्दु पूरा कऽ कऽ। संगहि ओकर एकटा शॉर्ट्स-संस्करण काटि कऽ अलगसँ अपलोड करू। एक सप्ताह बाद दुनूक आँकड़ा (व्यू, वॉच-टाइम) तुलना करू।

मैथिली सामग्रीक विशेष बात

ई अध्याय एहि पोथीक हृदय थिक। औजार सार्वभौम अछि, मुदा हमर प्रयोजन विशेष — मैथिलीमे, मिथिला लेल।

पहिल बात — भाषा-शुद्धता। ए.आइ. उपकरण सभ मैथिली लिखबामे प्रायः हिन्दीक छाया आनि दैत अछि — कारण ओ हिन्दीसँ बेसी सिखने अछि। चैटबॉटसँ बनाओल पटकथा, अनुवाद, उपशीर्षक — सभकेँ अपन आँखिसँ जाँचू। सोझ नियम: ए.आइ.क मैथिली प्रारूप (ड्राफ्ट) थिक, प्रमाण नहि। जतऽ शंका हो, अपन कान पर विश्वास करू — बाजि कऽ देखू, जे कान खटकए से हिन्दीक छाया।

दोसर — लिपिक प्रश्न। मैथिली देवनागरीमे लिखल जाइत अछि आ एकर अपन प्राचीन लिपि तिरहुता (मिथिलाक्षर) सेहो अछि। वीडियोक उपशीर्षक आ पाठ-पट देवनागरीमे राखू — सर्वाधिक लोक पढ़ि सकताह। तिरहुताक प्रयोग शोभा आ पहिचान लेल करू — शीर्षक-पटमे, लोगोमे, जलचिह्नमे। एहिसँ लिपिक दर्शनो होइत रहत आ जिज्ञासा सेहो जगत। (स्मरण रहए — ए.आइ. चित्र-उपकरण देवनागरी वा तिरहुता अक्षर शुद्ध नहि लिखि पबैत अछि; लिखल अक्षर सदिखन सम्पादनमे अपने जोड़ू, ए.आइ.सँ नहि लिखबाउ।)

तेसर — दृश्यक प्रामाणिकता। ए.आइ.केँ "भारतीय गाम" कहला पर ओ एकटा सामान्यीकृत उत्तर-भारतीय दृश्य बनाओत — जे मिथिला नहि थिक। प्रॉम्प्टमे मिथिलाक विशिष्ट संकेत दिअ: पोखरि, बरक गाछ, आमक गाछी, धानक खेत, माछ, पान, फूसक घर, आङनक तुलसी चौरा, भित्ति पर अरिपन-सन अलंकरण। तैयो जे बनत से "मिथिला-सन" हएत, मिथिला नहि — तेँ जतऽ सम्भव हो, असली फोटो-वीडियो मिलाउ (अध्याय ७ क विधिसँ) । ए.आइ.-दृश्य आ असली दृश्यक मिश्रण सभसँ प्रामाणिक परिणाम दैत अछि।

चारिम — मधुबनी/मिथिला चित्रकलाक मर्यादा। ए.आइ.केँ "मधुबनी शैलीमे" बनएबाक निर्देश देल जा सकैत अछि, आ ओ रंग-रेखाक नकल कऽ सेहो देत। मुदा एतऽ ठहरि कऽ सोचू। मिथिला चित्रकला जीवित परम्परा थिक — एकर आजीविका हजारो कलाकारक हाथमे अछि, आ प्रत्येक शैली (भरनी, कचनी, गोदना, गोबर) क अपन वंश-परम्परा अछि। ए.आइ.क बनाओल "मधुबनी-सन" चित्र ओहि परम्पराक श्रम आ ज्ञानक बिना ओकर रूप उठबैत अछि। हमर सुझाव: वीडियोमे असली कलाकारक असली कृति देखाउ (अनुमति आ श्रेय सहित) — ई कलाकारक सम्मानो थिक आ अहाँक वीडियोक शक्तियो। ए.आइ.क "मधुबनी-सन" अलंकरणक प्रयोग करी तँ स्पष्ट कही जे ई ए.आइ.-निर्मित अनुकरण थिक, प्रामाणिक मधुबनी नहि।

पाँचम — विषयक अक्षय भण्डार। मैथिली वीडियोक क्षेत्र एखनो लगभग खाली अछि — जे बनाओत, से पहिल हएत। किछु दिशा: बाल-सामग्री (गीत, कथा, अक्षर-ज्ञान — बाल-दर्शक सभसँ तेजीसँ बढ़ैत वर्ग थिक); पाबनि-तिहारक व्याख्या (छठि, सामा-चकेबा, जुड़ि-शीतल — की, किएक, केना); व्यंजन-विधि; लोककथा आ विद्यापति-पदक सचित्र प्रस्तुति; गाम-घरक शब्दावली (जे शब्द बिला रहल अछि तकर दृश्य-कोश); मिथिलाक स्थान-परिचय। प्रत्येक दिशा एक-एक चैनल भऽ सकैत अछि।

आ अन्तिम बात — गुणवत्ताक धैर्य। मैथिलीमे दर्शक-संख्या हिन्दीसँ कम हएत, ई स्वाभाविक। मुदा मैथिली दर्शकक निष्ठा बेसी होइत अछि — ओ टिप्पणी करताह, बँटताह, आ बेर-बेर घुरि कऽ औताह। संख्या नहि, सम्बन्ध देखू। सय निष्ठावान दर्शक दस हजार उदासीनसँ मूल्यवान।

अभ्यास: उपरि देल छह दिशामेसँ एकटा चुनू आ ओहि पर लगातार तीन वीडियोक योजना (विषय + एक-पंक्ति सार) बनाउ। तीन एहि द्वारे जे एकटा वीडियो प्रयोग थिक, तीनटा दिशा।

कॉपीराइट, नैतिकता आ जिम्मेदारी

शक्तिशाली औजारक संग जिम्मेदारी अबैत अछि। ई अध्याय छोट अछि, मुदा एकर प्रत्येक पाँती व्यवहारमे आनू।

कॉपीराइट — मूल सिद्धान्त: जे वस्तु अहाँ नहि बनौलहुँ, से बिना अनुमतिक प्रयोग नहि। फिल्मी गीत, आनक वीडियो-अंश, पोथीक पाठ, आनक फोटो — सभ पर ई लागू। "सभ कियो करैत अछि" कोनो तर्क नहि; यूट्यूबक स्वचालित प्रणाली (कन्टेन्ट आइ.डी.) पकड़ि लैत अछि आ दण्ड चैनलकेँ भोगऽ पड़ैत अछि। मुक्त-लाइसेन्स सामग्रीमे सेहो शर्त पढ़ू — केओ "श्रेय अनिवार्य" कहैत अछि, केओ "व्यावसायिक प्रयोग वर्जित"।

अपन बनाओल ए.आइ.-सामग्रीक अधिकारक प्रश्न सेहो जानू: अधिकांश उपकरणक शर्तमे बनाओल वीडियो-चित्रक व्यावसायिक प्रयोगक अनुमति सशुल्क योजनामे रहैत अछि, निःशुल्कमे कखनो सीमित। जाहि उपकरणसँ नियमित काज करी, तकर उपयोग-शर्त एक बेर अवश्य पढ़ू — विशेषतः "ओनरशिप" आ "कमर्शियल यूज" खण्ड।

व्यक्तिक गरिमा — तीन कठोर वर्जना: ककरो चेहरा-स्वरक अनुकरण बिना लिखित सहमतिक नहि (जीवित वा दिवंगत — दुनू); ककरो एहन दृश्यमे नहि देखाएब जे ओकर मान घटाबए वा ओकरा मुँहमे एहन बात नहि देब जे ओ कहियो नहि कहलक ("डीपफेक"); आ बालक-बालिकाक छवि-स्वरक विषयमे विशेष सतर्कता — अपन बच्चोक वीडियो राखबासँ पहिने सोचू जे ई सार्वजनिक भऽ रहल अछि।

सत्यक प्रति — ए.आइ.सँ यथार्थ-सन दृश्य बनि सकैत अछि, तेँ भ्रमक सम्भावना सेहो। नियम: कल्पना कल्पना कहाए, समाचार समाचार। ऐतिहासिक-सांस्कृतिक विषय पर वीडियो बनाबी तँ तथ्य अपन विश्वसनीय स्रोतसँ जाँचू — ए.आइ. चैटबॉट आत्मविश्वाससँ भूलो बजैत अछि (एकरा "हैलुसिनेशन" कहल जाइत अछि) । जतऽ ए.आइ.-दृश्य यथार्थ-सन हो ततऽ घोषणा करू (अध्याय १३) । दर्शकक विश्वास अर्जित करबामे बरख लगैत अछि, गमएबामे एक वीडियो।

आ अपना प्रति — ए.आइ.क सुविधा कखनो आलस्यमे नहि बदलए। ए.आइ.सँ काज लिअ, सोच नहि। जाहि दिन पटकथा, चयन, आ अन्तिम निर्णय ए.आइ.केँ सौंपि देब, ताहि दिन वीडियो "अहाँक" नहि रहत — आ दर्शक ई अन्तर सूँघि लैत छथि। औजार हाथमे राखू, हाथ औजार नहि बनए।

अभ्यास: अपन बनाओल पहिल वीडियोक "जिम्मेदारी-जाँच" करू — संगीतक लाइसेन्स? चित्र-स्रोतक अनुमति? तथ्यक जाँच? ए.आइ.-घोषणा? चारूक उत्तर लिखित राखू। ई अभ्यास प्रत्येक वीडियोक संग दोहराउ — किछुए दिनमे स्वभाव बनि जाएत।

अभ्यास-परियोजना: पहिल वीडियो आद्योपान्त

आब सभ सूत्र एक ठाम। ई अध्याय एकटा पूर्ण परियोजनाक नमूना थिक — "छठि पाबनि: सूर्य-उपासनाक महापर्व" नामक ९० सेकेण्डक वीडियो। एकरा पढ़ू, फेर अपन विषय पर एहीक्रमे दोहराउ।

दिन १ — योजना (अध्याय ३-४): उद्देश्य — छठिक मूल भावना आ विधि सरलतासँ बुझाएब। दर्शक — प्रवासी मैथिल परिवार, विशेषतः नव पीढ़ी। अवधि — ९० सेकेण्ड, अर्थात् नौ-दसटा क्लिप। चैटबॉटसँ दू-स्तम्भ पटकथा बनबाओल, अपने तीन ठाम सुधारल — "नहाय-खाय"क वर्णनमे स्थानीय विवरण जोड़ल, अन्तिम पाँतीमे भाव गहींर कएल। स्टोरीबोर्डमे दस दृश्य: भोरक नदी-घाट (स्थापना), ठेकुआ बनैत आङन, अर्घ्यक सूप-दउरा, साँझुक अर्घ्य, भोरुका अर्घ्य, पारण, आ समापन-पट (समापन) ।

दिन २ — स्वर (अध्याय ८): वॉयसओवर अपन स्वरमे रिकॉर्ड कएल — राति, शान्त कोठलीमे, अनुच्छेद-अनुच्छेद। सम्पादकक "एन्हान्स वॉयस"सँ निखारल। कुल ८२ सेकेण्ड — अर्थात् दृश्य-योजना ठीक बैसल।

दिन ३-४ — दृश्य (अध्याय ५-७): प्रत्येक दृश्यक छह-अंग प्रॉम्प्ट बनाओल (मैथिलीमे सोचल, अंग्रेजीमे अनुवाद) । मिथिला-संकेत प्रत्येकमे — "उत्तर बिहारक नदी-घाट, केराक थम्ह, बाँसक कोनिया, सूप-दउरामे फल"। तीन क्लिप पहिल प्रयासमे नीक भेटल; चारि क्लिपमे दू-तीन सुधार-चक्र लागल; दू दृश्य (ठेकुआ आ सूप) क लेल इमेज-टू-वीडियो विधि अपनाओल — पहिने चित्र सुधारल, फेर सूक्ष्म गति देल; आ एकटा दृश्यमे अपन खिचल असली फोटो प्रयोग कएल — घाटक भीड़, जे ए.आइ. कहियो ओहन नहि बना सकैत छल। (यैह मिश्रण अध्याय १४ क सार थिक।)

दिन ५ — संगीत आ जोड़ (अध्याय १०-११): ए.आइ.-संगीत — "धीमा, भक्तिभाव, बाँसुरी आ मृदंग, ९० सेकेण्ड"; लाइसेन्स जाँचल। सम्पादनक क्रम — पहिने स्वर, तखन क्लिप, सोझ कट, दू ठाम फेड; सभ क्लिप पर एक्के हल्लुक गरम फिल्टर; शीर्षक-पट पर तिरहुतामे "छठि" आ देवनागरीमे पूरा शीर्षक।

दिन ६ — उपशीर्षक आ प्रकाशन (अध्याय १२-१३): पटकथासँ मैथिली एस.आर.टी., फेर हिन्दी आ अंग्रेजी अनुवाद-फाइल (चैटबॉटसँ अनुवाद, अपने जाँच) । थम्बनेल — साँझुक अर्घ्यक दृश्य, तीन शब्द: "छठि किएक?"। अपलोडमे भाषा मैथिली, ए.आइ.-घोषणा सक्रिय, विवरणमे संगीत-श्रेय आ स्रोत। संगहि ३५ सेकेण्डक शॉर्ट्स-संस्करण।

दिन ७ — समीक्षा (अध्याय १५): जिम्मेदारी-जाँच चारू बिन्दु पर। परिवारकेँ देखाओल — सुझाव आएल जे भोरुका अर्घ्यक दृश्य छोट लगैत अछि; अगिला वीडियोक टीपमे लिखल।

देखू — सात दिन, प्रतिदिन एक-दू घण्टा, आ शून्यसँ प्रकाशित वीडियो धरि। पहिल परियोजनामे एहिसँ बेसियो समय लागत — लागए दियौ। तेसर-चारिम वीडियो धरि ई क्रम स्वभाव बनि जाएत आ समय आधा रहि जाएत।

अन्तिम बात। ई पोथी एतऽ सठैत अछि, मुदा सीखब एतऽसँ प्रारम्भ होइत अछि। उपकरण बदलत, मॉडल बदलत, स्क्रीन बदलत — मुदा जे ढाँचा अहाँ सीखलहुँ (योजना, प्रॉम्प्ट, सुधार-चक्र, जोड़, जिम्मेदारी) से बनल रहत। आब जाउ, आ मिथिलाक कथा संसारकेँ सुनाउ — अहींक स्वरमे, अहींक भाषामे।

शब्दावली

ए.आइ. (आर्टिफिशियल इंटेलिजेंस) — कृत्रिम बुद्धि; कम्प्यूटरक मनुष्य-सन सीखबा-बुझबाक क्षमता।

जेनरेटिव ए.आइ. — नव सामग्री (पाठ, चित्र, वीडियो, स्वर) रचनिहार ए.आइ.।

प्रॉम्प्ट — ए.आइ.केँ देल गेल लिखल निर्देश।

नकारात्मक प्रॉम्प्ट — जे नहि चाही तकर सूची।

टेक्स्ट-टू-वीडियो — लिखल वर्णनसँ वीडियो बनएबाक विधि।

इमेज-टू-वीडियो — स्थिर चित्रकेँ चलायमान करबाक विधि।

टी.टी.एस. (टेक्स्ट-टू-स्पीच) — लिखल पाठसँ बाजल स्वर बनएबाक विधि।

अवतार — स्क्रीन पर पाठ बजनिहार डिजिटल वक्ता।

लिप-सिंक — स्वर आ ठोरक गतिक मेल।

वॉयस क्लोनिङ — ककरो स्वरक डिजिटल प्रतिरूप बनाएब।

क्रेडिट — ए.आइ. उपकरणक प्रयोग-मुद्रा; प्रत्येक निर्माणमे किछु कटैत अछि।

मॉडल — ए.आइ.क कोनो विशेष संस्करण/इंजन।

क्लिप — वीडियोक छोट खण्ड (प्रायः ५-२० सेकेण्ड) ।

अनुपात (आस्पेक्ट रेशियो) — स्क्रीनक चौड़ाइ-ऊँचाइक अनुपात; 16:9 (चौड़ा), 9:16 (ठाढ़) ।

रिजॉल्यूशन — चित्रक सूक्ष्मता; 1080पी मानक, 4के अति-सूक्ष्म।

वॉटरमार्क — सामग्री पर छपल पहिचान-चिह्न।

स्टोरीबोर्ड — दृश्य-दर-दृश्य खाका।

बी-रोल — मुख्य वक्ताक अतिरिक्त सहायक दृश्य।

टाइमलाइन — सम्पादकमे समयक रेखा जाहि पर क्लिप सजैत अछि।

ट्रैक (पट्टी) — टाइमलाइनक तह — वीडियो, स्वर, संगीत आदि।

कट — क्लिप काटब वा एक क्लिपसँ दोसर पर सोझे जाएब।

ट्रांजिशन (संक्रमण) — दू क्लिपक बीचक जोड़-ढंग (फेड आदि) ।

फेड इन/आउट — धीरे-धीरे उभरब/बिलाएब।

रेंडर/निर्यात (एक्सपोर्ट) — सम्पादित वीडियोकेँ अन्तिम फाइल बनाएब।

एस.आर.टी. — उपशीर्षकक मानक फाइल-प्रारूप।

बर्न्ड-इन — वीडियोमे स्थायी रूपेँ सटल उपशीर्षक।

थम्बनेल — वीडियोक मुख-चित्र।

हैलुसिनेशन — ए.आइ.क आत्मविश्वासपूर्ण भूल।

कन्टेन्ट आइ.डी. — यूट्यूबक कॉपीराइट-जाँच प्रणाली।

डीपफेक — ककरो चेहरा-स्वरक भ्रामक अनुकरण।

मुद्रीकरण (मोनेटाइजेशन) — वीडियोसँ आय।

प्रॉम्प्ट-संग्रह

नीचाँ दसटा तैयार प्रॉम्प्ट-ढाँचा अछि — मैथिलीमे। कोष्ठकक स्थान अपन विवरणसँ भरू आ अन्तिम प्रॉम्प्ट अंग्रेजीमे बनाउ (अध्याय ५ क विधिसँ) ।

१. गामक भोर: "उत्तर बिहारक गाम, भोर, खेत पर हल्लुक कुहेस, फूसक घर, दूरमे बरक गाछ, [अपन विवरण], कैमरा धीरे दहिना घूमैत, नरम सुनहरा प्रकाश, यथार्थवादी वृत्तचित्र शैली।"

२. पोखरि-दृश्य: "कमल-पुरइनसँ भरल पोखरि, [ऋतु], कात पर [गाछ/घाट], पानिमे हल्लुक लहरि, [पक्षी/माछ]क सूक्ष्म गति, कैमरा अत्यन्त धीरे आगाँ, [समय]क प्रकाश, सिनेमाई।"

३. आङन-दृश्य: "मिथिलाक घरक आङन, तुलसी चौरा, [क्रिया — जेना चूल्हि पर भोजन बनैत], भित्ति पर लोक-अलंकरण, गरम आत्मीय प्रकाश, क्लोज-अप, यथार्थवादी।"

४. पाबनि-दृश्य: "[पाबनिक नाम]क तैयारी, [मुख्य वस्तु — जेना सूप-दउरा, दीप], आङन/घाटक दृश्य, व्यस्त हाथ सभक क्लोज-अप, उत्सवक उल्लास, गरम प्रकाश, वृत्तचित्र शैली।"

५. प्रकृति-दृश्य: "धानक हरियर खेत हवामे लहराइत, [समय], दूर क्षितिज, [पक्षी] उड़ैत, ड्रोन दृश्य धीरे ऊपर उठैत, प्राकृतिक प्रकाश, अति-सूक्ष्म विवरण।"

६. वस्तु-परिचय (व्यंजन/शिल्प): "[वस्तु]क अत्यन्त लग्गसँ दृश्य, [पृष्ठभूमि], हल्लुक भाप/चमक, कैमरा धीरे-धीरे वस्तुक चारू कात घूमैत, स्टूडियो-सन नरम प्रकाश, स्वादिष्ट/आकर्षक शैली।"

७. कथा-चित्रण (एनीमेशन): "[पात्र-वर्णन] [क्रिया करैत], [स्थान], हाथसँ बनाओल चित्र-सन एनीमेशन शैली, कोमल जलरंग, बाल-पोथीक चित्र-सन, धीमा सुखद गति।"

८. ऐतिहासिक कल्पना: "[काल]क मिथिलाक कल्पित दृश्य, [क्रिया], पुरान ताड़पत्र-चित्र-सन रंग, धीमा गम्भीर कैमरा, संग्रहालय-वृत्तचित्र शैली।" (संग ए.आइ.-घोषणा अनिवार्य — ई कल्पना थिक।)

९. शीर्षक-पट पृष्ठभूमि: "अमूर्त अलंकरण-पैटर्न धीरे-धीरे चलैत, [रंग-योजना — जेना सिन्दूरिया लाल आ पीत], कोनो अक्षर नहि, कोनो आकृति नहि, शान्त एकरस गति, 10 सेकेण्ड लूप।" (अक्षर सम्पादनमे अपने जोड़ू।)

१०. गति-निर्देश (इमेज-टू-वीडियो लेल): "एहि चित्रमे [अंग — जेना पानि/पात/वस्त्र]मे हल्लुक स्वाभाविक गति, कैमरा [स्थिर/अत्यन्त धीरे आगाँ], आर किछु नहि बदलए, चित्रक मूल रंग-रेखा अक्षुण्ण रहए।"

उपकरण-दिशा (२०२६ क स्थिति)

ई सूची लिखबा कालक (२०२६) स्थिति थिक; नाम आ सुविधा बदलैत रहैत अछि — श्रेणी बुझू, नाम नहि रटू। कोनो उपकरण अपनएबासँ पहिने तीन बात जाँचू: निःशुल्क योजनामे की भेटैत अछि, बनल सामग्रीक व्यावसायिक प्रयोगक अनुमति, आ वॉटरमार्कक स्थिति।

पाठ-सँ-वीडियो / चित्र-सँ-वीडियो: गूगल वीओ (Veo), ओपनएआइ सोरा (Sora), क्लिङ (Kling), रनवे (Runway), पिक्सवर्स (PixVerse), सीडान्स (Seedance), हैलुओ (Hailuo), लूमा (Luma), पिका (Pika) । — यथार्थवादी दृश्य, सिनेमाई क्लिप।

अवतार वीडियो: हेजेन (HeyGen), सिन्थेसिया (Synthesia) । — डिजिटल वक्ता, लिप-सिंक, बहुभाषी।

पाठ-सँ-स्वर आ स्वर-अनुकरण: इलेवनलैब्स (ElevenLabs) आदि; भारतीय भाषा-केन्द्रित सेवा सभ सेहो बढ़ि रहल अछि — मैथिली-विकल्प एहीमे पहिने आओत, ताकैत रहू।

ए.आइ. संगीत: सूनो (Suno), उडियो (Udio) आदि — वर्णन वा बोलसँ संगीत/गीत।

चित्र-निर्माण: प्रायः प्रत्येक वीडियो-उपकरणहिमे सम्मिलित; अलगसँ मिडजर्नी (Midjourney), आ चैटबॉट-सम्बद्ध चित्र-सुविधा सभ।

सम्पादन: कैपकट (CapCut) — मोबाइल-कम्प्यूटर, सरल, ऑटो-कैप्शन सहित; डा-विन्ची रिजॉल्व (DaVinci Resolve) — कम्प्यूटर, निःशुल्क, व्यावसायिक स्तर; कैनवा (Canva) — थम्बनेल, बैनर, सरल वीडियो।

एक-मे-सभ (योजना-सँ-वीडियो): इनवीडियो (InVideo)-सन उपकरण, जे विषय देला पर पटकथा-दृश्य-स्वर सभ स्वयं जोड़ि दैत अछि — द्रुत काज लेल; मुदा एहिमे अहाँक अपन छाप कम रहैत अछि, तेँ सीखबा लेल एहि पोथीक डेग-दर-डेग विधिए नीक।

चैटबॉट (पटकथा, अनुवाद, एस.आर.टी., विचार-मन्थन): क्लॉड (Claude), चैटजीपीटी (ChatGPT), जेमिनी (Gemini) ।

आ सभसँ अन्तमे — ई पोथियो कालक्रमे पुरान हएत, मुदा अहाँ नहि। जे "स्वयं सीखू" स्वभाव एहि पोथीसँ बनल, से प्रत्येक नव उपकरण पर लागत। नव उपकरण देखी तँ तीन प्रश्न पूछू — प्रॉम्प्ट कतऽ लिखाइत अछि? सेटिङमे की-की अछि? परिणाम कतऽ आ केना अबैत अछि? — आ अहाँ पाँचे मिनटमे ओकरा चला लेब।

भाग ३
उन्नत तकनीकी अध्ययन

कृत्रिम बुद्धि (ए.आइ.) वर्तमान स्थितिमे प्रविधिक दुनियामे एकटा क्रान्ति आनि देलक अछि, विशेष रूपसँ वीडियो निर्माणक क्षेत्रमे। पहिने जतय उच्च गुणवत्ताक वीडियो बनएबाक लेल महग कैमरा, पेशेवर सम्पादक, मोशन डिजाइनर आ पोस्ट-प्रोडक्शनमे घण्टों समयक आवश्यकता होइत छल, ओतय 2026 धरि ई प्रक्रिया मात्र किछु मिनटोंक खेल बनि गेल अछि। वर्तमान स्थितिमे, निर्माता, एजेन्सी आ व्यावसायिक प्रतिष्ठान मात्र एकटा साधारण टेक्स्ट प्रॉम्प्ट (पाठ्य निर्देशिका) क माध्यमसँ पेशेवर वीडियो उत्पन्न कए सकैत छथि, दृश्य सम्पादित कए सकैत छथि, ए.आइ. अवतार बना सकैत छथि आ स्वर (Voiceover) उत्पन्न कए सकैत छथि। ई पोथी 'स्वयं सीखू' (Teach Yourself) श्रृंखलाक अन्तर्गत ए.आइ. सँ मैथिली वीडियो निर्माणक सम्पूर्ण प्रक्रिया, उपकरण, प्रॉम्प्ट इंजीनियरिंग, स्वर-अनुकरण आ मूल्यांकनक विस्तृत विश्लेषण प्रस्तुत करैत अछि। एहि पोथीमे व्यावहारिक उदाहरण आ अन्तरफलकक विवरणात्मक स्क्रीनशॉट शामिल कएल गेल अछि, जाहिसँ पाठक कैं कोनो ए.आइ. औजारक उपयोग करबामे पूर्ण स्पष्टता भेट सकै।

2026 क प्रमुख ए.आइ. वीडियो जेनरेटर उपकरण

ए.आइ. वीडियो जेनरेटर प्लेटफार्मक चयन करब एकटा जटिल काज भऽ सकैत अछि, किएक तँ प्रत्येक उपकरणक अपन विशिष्टता अछि। किछु उपकरण यथार्थवादी अवतार बनएबामे निपुण अछि, तँ किछु सिनेमाटिक दृश्य वा उन्नत सम्पादनमे। कोनो एक उपकरण पूर्ण नहि अछि, मुदा विभिन्न उपकरणक संयोजनसँ उत्कृष्ट परिणाम प्राप्त कएल जा सकैत अछि।

गूगल वीओ (Google Veo) आ एकर उन्नत संस्करण

गूगल वीओ एकटा अत्यधिक विश्वसनीय प्लेटफार्म अछि जे सिनेमाटिक ए.आइ. वीडियो आ उच्च-गुणवत्ता वला दृश्य कहानी कहबाक लेल जानल जाइत अछि। गूगल डीपमाइंड (Google DeepMind) द्वारा विकसित ई वीडियो-जेनरेशन मॉडल टेक्स्ट, चित्र वा वीडियोक सन्दर्भसँ नव वीडियो उत्पन्न करबामे सक्षम अछि।

मई 2024 मे अपन प्रारम्भिक घोषणाक बादसँ, गूगल वीओ 2, वीओ 3 आ नव वीओ 3.1 संस्करण जारी केलक अछि। गूगल वीओ 3.1 वर्तमानमे मोबाइल-अनुकूलित अछि आ एहिमे वर्टिकल वीडियो एकीकरण, उच्च रिज़ॉल्यूशन (4k धरि) आ एकाधिक सन्दर्भ चित्र (अधिकतम तीन) अपलोड करबाक सुविधा अछि। एहि उन्नत संस्करणमे दृश्य विस्तार (Scene extension) क सुविधा अछि, जाहिसँ पहिल क्लिप कैं लम्बा वीडियोमे बदलल जा सकैत अछि, आ संगे प्रथम तथा अन्तिम फ्रेम निर्दिष्ट करबाक नियन्त्रण सेहो देल गेल अछि।

चित्र-विवरण 1: गूगल वीओ 3.1 क डेशबोर्ड, जतय 'Upload Reference Images' क विकल्प अछि आ संगे आस्पेक्ट रेशियो चुनबाक ड्रॉपडाउन मेनू देखा रहल अछि।

विभिन्न संस्करणक मध्यक तकनीकी विकास कैं नीचाँ देल गेल सारणीमे स्पष्ट कएल गेल अछि:

फीचर (Feature)	वीओ 2 (Veo 2)	वीओ 3 (Veo 3)	वीओ 3.1 (Veo 3.1)
ऑडियो (Audio)	कोनो नहि, मात्र मूक वीडियो	हँ, उपलब्ध अछि	हँ, उन्नत साउण्डस्केप आ सिंक क संग समृद्ध
सन्दर्भ (References)	नहि	नहि	हँ, अधिकतम तीन सन्दर्भ चित्र
वीडियो अवधि	पाँच सँ आठ सेकण्ड	आठ सेकण्ड	चारि, छह, आ आठ सेकण्ड
रिज़ॉल्यूशन (Resolution)	720p	720p आ 1080p	720p, 1080p, 4k
नियन्त्रण सुविधा	अस्पष्ट प्रॉम्प्ट पालन; प्रकाशमे सीमित भिन्नता	उन्नत प्रॉम्प्ट पालन, कैमरा गति	प्रथम आ अन्तिम फ्रेम ट्रान्जिशन, दृश्य नियन्त्रण
गुणवत्ता (Quality)	नीक	बेहतर; भौतिकी आ कैमरा गतिमे सुधार	उत्कृष्ट; निरन्तरता उपकरण आ ऑडियोक समावेश

ए.आइ. वीडियो औजारक योजना, क्रेडिट, रेजोल्यूशन, वॉटरमार्क आ मूल्य बारम्बार बदलैत रहैत अछि। तेँ सीखबा काल छोट परीक्षणसँ प्रारम्भ करू, अपन आवश्यकतानुसार एक-दू औजार चुनू, आ स्थायी कौशल — प्रॉम्प्ट, दृश्य-पट्ट, कैमरा-भाषा, ध्वनि आ सम्पादन — पर बेसी ध्यान दिअ।

रनवे (Runway Gen 4.5) आ सृजनात्मक नियन्त्रण

रनवे (Runway) फिल्म निर्माता आ वीएफएक्स (VFX) कलाकार लोकनिक लेल एकटा शक्तिशाली उपकरण अछि। ई उपकरण ओहि लोकनिक लेल अछि जे मात्र प्रॉम्प्ट लिखि कें सन्तुष्ट नहि होइत छथि, बरु सम्पूर्ण सृजनात्मक प्रक्रिया पर नियन्त्रण चाहैत छथि। रनवे Gen 4.5 उन्नत कैमरा नियन्त्रण (पैन, टिल्ट, ज़ूम) आ 'मल्टी-मोशन ब्रश' जेहन अद्वितीय सुविधा प्रदान करैत अछि, जाहिसँ स्थिर चित्रक विशिष्ट हिस्सा कें एनिमेट कएल जा सकैत अछि।

एकर अन्तरफलक नव प्रयोक्ताक लेल किछु जटिल भऽ सकैत अछि। एकर साइडबारमे औजार, ऐप, चैट, कार्यप्रवाह आ लाइव जेहन विकल्प भरल रहैत अछि, जाहिसँ प्रॉम्प्ट बार खोजब प्रारम्भमे कठिन भऽ सकैत अछि। अध्ययनसँ ई सेहो ज्ञात होइत अछि जे यद्यपि एकर पृष्ठभूमि जनरेशन नीक होइत अछि, मुदा कहियो-काल चरित्रक गति अस्वाभाविक बुझाइत अछि आ चेहरा वा आँखिमे 'ग्लिच' (तकनीकी त्रुटि) देखल जाइत अछि।

चित्र-विवरण 2: रनवे Gen 4.5 क सम्पादन विन्डो, जतय 'मल्टी-मोशन ब्रश' औजारसँ चित्रक एकटा निश्चित भाग (जेना नदीक पानि) कें रंगल जा रहल अछि जाहिसँ मात्र ओतय गति उत्पन्न भऽ सकै।

सीडांस 2.5 (Seedance 2.5) क तीव्र गति

बाइटडांस (ByteDance) द्वारा विकसित सीडांस ए.आइ. वीडियो समुदायमे अपन गतिक लेल प्रसिद्ध अछि। Seedance 2.0 मात्र 30 सेकण्डमे 10 सेकण्डक क्लिप उत्पन्न कए सकैत छल, आ एकर 2.5 संस्करण (MuAPI) आ अधिक उन्नत अछि। ई 4 सँ 30 सेकण्ड धरिक वीडियो जनरेट करबामे सक्षम अछि आ जेपीईजी (JPEG), पीएनजी (PNG), वेबपी (WebP) सँ लऽ कें एमपी4 (MP4) आ एमओवी (MOV) धरिक इनपुट स्वीकार करैत अछि।

सीडांस 2.5 मे आस्पेक्ट रेशियोक विस्तृत विकल्प उपलब्ध अछि (जेना 16:9, 9:16, 1:1, 21:9), जाहिसँ ई सोशल मीडिया आ सिनेमाटिक दुनू प्रकारक वीडियोक लेल उपयोगी अछि। आउटपुट स्वरूपक लेल, ई डिफ़ॉल्ट रूपसँ H.264 (yuv420p) क संग MP4 जनरेट करैत अछि, मुदा उच्च रंग निष्ठा (higher color fidelity) आ बार-बार सम्पादनक लेल MOV (yuv444p) क विकल्प सेहो दैत अछि।

अन्य प्रमुख वीडियो जेनरेशन प्लेटफार्म

बाजारमे उपलब्ध अन्य प्लेटफार्म सेहो विशिष्ट उपयोगक लेल अपन अलग पहचान रखैत छथि।

HeyGen (हेजेन): ई उपकरण एकटा सम्पूर्ण प्रोडक्शन स्टूडियोक काज करैत अछि। ई प्रस्तुतकर्ता-आधारित (presenter-led) वीडियो उत्पन्न करैत अछि जाहिमे अवतार 175 सँ अधिक भाषामे सटीक लिप-सिंक (होंठक गति) क संग संवाद करैत अछि।

VidCap आ DeeVid AI: VidCap तीव्र आ सटीक कैप्शन (उपशीर्षक) आ बहुभाषी टेक्स्ट ओभरलेक लेल सर्वोत्तम मानल जाइत अछि। दोसर दिस, DeeVid AI चित्र वा टेक्स्ट कें छोट सोशल मीडिया वीडियोमे बदलबाक लेल सबसँ तेज विकल्प अछि।

LTX Studio: ई उपकरण 'शॉट-बाय-शॉट' स्टोरीबोर्डिंगक संग अत्यधिक रचनात्मक नियन्त्रण दैत अछि, जाहिसँ पूरा फिल्मक दृश्य-दर-दृश्य योजना बनायल जा सकैत अछि।

KlingAI आ Sora2: KlingAI अत्यन्त यथार्थवादी दृश्य उत्पन्न करैत अछि, मुदा एहिमे कहियो-काल तकनीकी त्रुटि (glitches) आबि जाइत अछि। ओतहि, ओपन-ए.आइ. (OpenAI) क Sora2 कहानी-आधारित क्लिप, ऑडियो आ लिप-सिंकक लेल बहुत शक्तिशाली अछि, यद्यपि एकर आउटपुट गुणवत्ता कहियो-काल असंगत भऽ सकैत अछि।

2026 क प्रमुख ए.आइ. वीडियो जेनरेटरक समग्र तुलना नीचाँ देल गेल अछि:

श्रेणी (Category)	सर्वोत्तम उपयोग (Best For)	योजना सम्बन्धी टिप्पणी
LTX Studio	शॉट-बाय-शॉट स्टोरीबोर्डिंगक संग नियन्त्रण	नि:शुल्क/सशुल्क योजना आ मूल्य समयानुसार बदलैत अछि।
Adobe Firefly	व्यावसायिक रूपसँ सुरक्षित ए.आइ. वीडियो	नि:शुल्क/सशुल्क योजना आ मूल्य समयानुसार बदलैत अछि।
Descript	स्क्रिप्ट सम्पादित कऽ कँ वीडियो सम्पादन	नि:शुल्क/सशुल्क योजना आ मूल्य समयानुसार बदलैत अछि।
VEED	तीव्र सामग्री उत्पादन आ पुनर्उपयोग	नि:शुल्क/सशुल्क योजना आ मूल्य समयानुसार बदलैत अछि।
InVideo AI	सोशल मीडिया लेल प्रॉम्प्ट-सँ-वीडियो	नि:शुल्क/सशुल्क योजना आ मूल्य समयानुसार बदलैत अछि।
OpusClip	लम्बा वीडियोसँ वायरल क्लिप निकालब	नि:शुल्क/सशुल्क योजना आ मूल्य समयानुसार बदलैत अछि।

चित्र-विवरण 3: HeyGen क अवतार चयन स्क्रीन, जतय विभिन्न आयु, लिंग आ परिधान वला यथार्थवादी डिजिटल मानव अवतारक सूची प्रदर्शित अछि, जकरा माउससँ क्लिक कऽ चुनल जा सकैत अछि।

प्रॉम्प्ट इंजीनियरिंगक सिद्धान्त आ विज्ञान (Prompt Engineering)

ए.आइ. वीडियो जेनरेटरसँ उत्कृष्ट परिणाम प्राप्त करबाक लेल 'प्रॉम्प्ट इंजीनियरिंग' (निर्देशिका निर्माण) क ज्ञान आवश्यक अछि। एकटा नीक प्रॉम्प्ट आदेशात्मक निर्देशसँ अधिक एकटा वर्णनात्मक चित्रक समान होइत अछि। केवल "एकटा सुन्दर सूर्यास्त बनाउ" लिखबाक बदलामे, यदि "पहाड़ पर एकटा जीवन्त सूर्यास्त, जतय गुलाबी बादलक बीचसँ सुनहला किरण आबि रहल अछि, लो-एंगलसँ लेल गेल दृश्य" लिखल जाए, तँ ए.आइ. ओकरा नीक जँका समझि सकैत अछि।

टेक्स्ट-टू-वीडियो प्रॉम्प्टक सूत्र

प्रॉम्प्टक निर्माण लेल एकटा मानकीकृत सूत्रक प्रयोग कएल जाइत अछि, जाहिसँ कोनो महत्वपूर्ण विवरण छूटि नहि जाए: प्रॉम्प्ट = विषय (Subject) + क्रिया (Action) + सेटिंग (Setting) + कैमरा (Camera) + प्रकाश (Lighting) + शैली (Style) + प्रारूप (Format) + प्रतिबन्ध (Constraints)।

ई तत्व कोना काज करैत अछि, एकर विस्तृत विवरण एहि प्रकार अछि:

विषय (Subject): ई वीडियोक मुख्य केन्द्र अछि। यदि मॉडल मुख्य विषय कँ तुरन्त नहि पहिचानि सकत, तँ आउटपुट अस्त-व्यस्त भऽ जाइत अछि। ई स्पष्ट आ विशिष्ट हेबाक चाही (जेना- "एकटा मैट ब्लैक स्किनकेयर बोतल", "एकटा युवा बरिस्ता", "काँचक परफ्यूम बोतल")। अमूर्त शब्द जेना "एकटा सुन्दर जीवन शैली", "एकटा ब्राण्ड" वा "सुन्दर दृश्य" क प्रयोग मुख्य विषयक रूपमे नहि करबाक चाही।

क्रिया (Action): वीडियोमे गति आवश्यक अछि। बिना गतिक प्रॉम्प्ट एकटा स्थिर चित्र जेहन लागैत अछि। विषय कँ एकटा मुख्य दृश्यमान क्रिया देल जेबाक चाही (जेना- "परफ्यूम बोतल धीरे-धीरे घूमि रहल अछि जखन कि पाछाँसँ कोमल धुआँ उठि रहल अछि")। एकटा छोट क्लिपमे पाँच अलग-अलग क्रियाक निर्देश देबासँ बचबाक चाही; आवश्यकता पड़ला पर अलग-अलग क्लिप बना कऽ सम्पादनमे जोड़ल जाए।

सेटिंग (Setting): ई ए.आइ. मॉडल कँ दृश्य कँ बुझबामे सन्दर्भ प्रदान करैत अछि। एहिमे स्थान, दिनक समय, मौसम, पृष्ठभूमि विवरण आ सतहक सामग्री शामिल अछि। (जेना- "रातिक समय एकटा छोट स्वतन्त्र किताबक दोकानमे, जतय काठक शेल्फ अछि आ खिरकी पर पानि पड़ि रहल अछि")।

कैमरा (Camera): शॉट आकार, कोण आ गति (जेना- क्लोज-अप, स्लो पुश-इन, शैलो डेप्थ ऑफ फील्ड) ।

प्रकाश (Lighting): मूड, दिनक समय आ प्रकाशक स्रोत (जेना- वार्म मॉर्निंग लाइट, सॉफ्ट रिफ्लेक्शन) ।

शैली (Style): दृश्य रूप वा विधा (जेना- यथार्थवादी कमर्शियल प्रोडक्ट वीडियो, 3D एनीमेशन) ।

प्रारूप (Format): अवधि, पहलू अनुपात (aspect ratio), आ गति (जेना- इंस्टाग्राम रील्स लेल 6-सेकण्डक वर्टिकल वीडियो) ।

प्रतिबन्ध (Constraints): ई ओ चिज अछि जकरा वीडियोमे नहि रखबाक अछि। नकारात्मक प्रॉम्प्ट जेना "नो टेक्स्ट, नो लोगो, नो ग्लिचिंग" आदि।

चित्र-विवरण 4: InVideo AI क प्रॉम्प्ट बॉक्स, जाहिमे उपरि देल सूत्र (Subject + Action + Setting + Camera) क उपयोग कऽ एकटा विस्तृत अंग्रेजी प्रॉम्प्ट लिखल गेल अछि आ ओकर नीचाँ 'Generate' बटन अछि।

मल्टी-शॉट जनरेशनक सूत्र (Multi-shot Formula)

जखन कोनो पैघ कहानी कहबाक हो, तँ 'मल्टी-शॉट फॉर्मूला' (जेना Wan 2.7/2.6 मॉडल लेल) क उपयोग कएल जाइत अछि। मल्टी-शॉट प्रॉम्प्ट = समग्र विवरण (Overall description) + शॉट संख्या (Shot number) + टाइमस्टैम्प (Timestamp) + शॉट सामग्री (Shot content) ।

एहिमे सम्पूर्ण वीडियोक एकटा छोट सारांश प्रारम्भमे देल जाइत अछि। प्रत्येक शोटमे सन्दर्भ पहचानकर्ता (Reference identifier जेना 'Image 1' वा 'character1') क उपयोग कएल जेबाक चाही जाहिसँ अलग-अलग क्लिपमे पात्रक पहचान (identity) निरन्तर बनल रहय। संवाद (Lines) कैँ स्पष्ट रूपसँ लिखल जेबाक चाही, जेना 'X says: "xxx"', जहिसँ मॉडल संवाद कैँ ज्यों-का-त्यों सुरक्षित रखत। यदि संवाद नहि चाही तँ 'No dialogue' क स्पष्ट निर्देश देल जाए।

कैमराक कोण, गति आ प्रकाश व्यवस्था

सिनेमाटोग्राफीक भाषा कैँ ए.आइ. वीडियो जेनरेटर लग पहुँचाएब एकटा विशिष्ट कला अछि। सही शब्दावलीक उपयोग वीडियो कैँ शौकिया (amateur) सँ पूर्ण पेशेवर बना सकैत अछि।

कैमरा कोण (Camera Angles)

आई-लेवल शॉट (Eye-level shot): ई एकटा तटस्थ परिप्रेक्ष्य प्रदान करैत अछि, जेना कोनो मानव अपन ऊँचाईसँ देखि रहल हो (उदा. "चाय पिबैत महिलाक आई-लेवल शॉट") ।

लो-एंगल शॉट (Low-angle shot): एहिमे कैमरा विषयक नीचाँ रहैत अछि आ ऊपरक दिस देखैत अछि, जकरासँ विषय शक्तिशाली वा प्रभावशाली बुझाइत अछि (उदा. "सुपरहीरोक लो-एंगल ट्रैकिंग शॉट") ।

हाई-एंगल शॉट (High-angle shot): कैमरा विषयक ऊपर रहैत अछि, जकरासँ विषय छोट, कमजोर वा भीड़ क हिस्सा प्रतीत होइत अछि।

बर्ड्स-आई व्यू (Bird's-eye view): सीधा ऊपरसँ लेल गेल शॉट, जे सम्पूर्ण दृश्यक नक्शा (map-like) जेहन परिप्रेक्ष्य दैत अछि।

वर्म्स-आई व्यू (Worm's-eye view): जमीनसँ सीधा ऊपरक दिस देखैत अत्यंत नीचाँ कोणक शॉट, जे गगनचुम्बी इमारत आदिक विशालता देखाबय लेल प्रयुक्त होइत अछि।

डच एंगल (Dutch angle): कैमरा कैँ एक दिस झुका देल जाइत अछि (canted angle), जकरासँ बेचैनी, भटकाव वा गतिशीलताक बोध होइत अछि।

कैमरा गति (Camera Movements)

प्रत्येक कैमरा गति एकटा भावनात्मक उद्देश्य कैं पूरा करैत अछि।

स्टेटिक शॉट (Static shot): कैमरा बिल्कुल स्थिर रहैत अछि, जकरासँ शांति आ ध्यानाक्रेन्द्रित हेबाक संकेत भेटैत अछि।

पैन (Pan left/right): कैमरा अपन अक्ष पर क्षैतिज रूपसँ बामा वा दहिना घूमैत अछि। ई दर्शक कैं दृश्यसँ परिचित कराबय लेल उपयोगी अछि।

टिल्ट (Tilt up/down): कैमरा अपन अक्ष पर ऊर्ध्वाधर रूपसँ ऊपर वा नीचाँ घूमैत अछि।

डॉली (Dolly in/out): कैमरा शारीरिक रूपसँ विषयक लग वा दूर जाइत अछि। 'पुश-इन' घनिष्ठता वा तनाव पैदा करैत अछि, जखन कि 'पुल-आउट' अलगाव वा पैमाना कैं दर्शाबैत अछि।

ट्रक (Truck left/right): कैमरा क्षैतिज रूपसँ विषयक समानान्तर चलैत अछि, प्रायः चलैत पात्रक संग-संग।

पेडस्टल (Pedestal up/down): कैमरा अपन स्तर (level) बनौने राखि कैं ऊपर वा नीचाँ उठैत अछि (जेना विशाल गाछ कैं देखाबय लेल) ।

ज़ूम (Zoom in/out): कैमराक लेन्स अपन फोकल लम्बाई बदलैत अछि। ई डॉलीसँ भिन्न अछि किएक तँ एहिमे कैमरा स्वयं नहि घूमि रहल अछि।

क्रेन शॉट (Crane shot): विस्तृत प्रकटीकरण (dramatic reveals) वा पैघ युद्ध मैदान देखाबय लेल एकटा क्रेन पर लागल कैमराक उपयोग।

विप पैन (Whip pan): अत्यन्त तीव्र पैन जकरासँ चित्र धुंधला भऽ जाइत अछि, प्रायः ट्रांजिशनक लेल वा तीव्र गति देखाबय लेल प्रयुक्त होइत अछि।

आर्क शॉट (Arc shot / Orbit): कैमरा विषयक चारू दिस एकटा वृत्ताकार पथमे घूमैत अछि, जे विषयक महत्व कैं दर्शाबैत अछि।

चित्र-विवरण 5: कैमरा गतिक दृष्टान्त, जतय 'पैन', 'टिल्ट' आ 'डॉली' कैं तीरक निशान (arrows) सँ एकटा 3D मॉडलक चारू दिस देखाओल गेल अछि, जाहिसँ पाठक कैं गतिक दिशा स्पष्ट भऽ सकै।

प्रकाश व्यवस्था (Lighting Parameters)

प्रकाश व्यवस्था वीडियोक गुणवत्ता आ मूड निर्धारित करबामे सर्वोपरि अछि। सांख्यिकीय रूपसँ, प्रकाशक वर्णन वीडियोक गुणवत्ता पर सबसँ पैघ प्रभाव डालैत अछि।

सिनेमाटिक वार्म लाइट: 'Golden hour cinematography' क उपयोग कऽ कैं भोर वा साँझक सुनहला प्रकाश प्राप्त कएल जाइत अछि।

चियारोस्चुरो (Chiaroscuro): नाटक वा रहस्यमय दृश्य लेल गहरा छाया आ प्रकाशक उच्च कन्ट्रास्ट।

थ्री-पॉइन्ट लाइटिंग (Three-point lighting): 'की लाइट' (मुख्य प्रकाश जे रंग आ दिशा तय करैत अछि), 'फिल लाइट' (छाया कम करबा लेल), आ 'बैकलाइट/रिम लाइट' (विषय कैं पृष्ठभूमि सँ अलग करबा लेल, जेना केश वा कान्हक चमक) क सन्तुलित संयोजन।

प्रॉम्प्टमे "ड्रामैटिक लाइटिंग" (dramatic lighting) लिखबाक बदलामे, विशिष्ट स्रोत निर्दिष्ट करब अधिक लाभदायक होइत अछि, जेना "अवतारक चेहरा पर स्क्रीनक बामा दिससँ एकटा गरम (warm) की-लाइट, आ पाछाँसँ हल्का उज्जर बैकलाइट जे ओकरा पृष्ठभूमिसँ अलग कए रहल अछि"।

ऑडियो, टेक्स्ट-टू-स्पीच आ स्वर-अनुकरण

कोनो उच्च-गुणवत्ता वला वीडियो बिना उत्कृष्ट श्रव्य (Audio) कैं पूर्ण नहि मानल जाइत अछि। ए.आइ. ऑडियो तकनीक आब अत्यन्त उन्नत भऽ चुकल अछि, जाहिमे इलेवनलैब्स (ElevenLabs) जेहन प्लेटफार्म अग्रणी अछि।

टेक्स्ट-टू-स्पीच (TTS) मॉडल आ इलेवनलैब्स

इलेवनलैब्स ए.आइ. वॉइस जेनरेटर आब 70 सँ अधिक भाषामे पाठ्यसँ ध्वनि (Text-to-Speech) उत्पन्न कए सकैत अछि, जाहिमे हिन्दी आ भारतीय क्षेत्रीय लहजा (Indic/Desi accents) सेहो शामिल अछि। यद्यपि 2026 धरि मैथिली लेल सीधा पूर्व-प्रशिक्षित 'मैथिली' ट्याग वला मॉडल सर्वसुलभ नहि भऽ सकैत अछि, मुदा हिन्दी वा 'देसी' (Desi) लहजाक उपयोग कए वा कस्टम स्वर-अनुकरणक माध्यमसँ मैथिली वीडियो लेल पूर्ण रूपसँ यथार्थवादी ध्वनि उत्पन्न कएल जा सकैत अछि। देवनागरी लिपिमे लिखल मैथिली पाठ्य कें हिन्दी (Desi) मॉडल बहुत स्पष्टतासँ बाजि सकैत अछि।

इलेवनलैब्स विभिन्न आवश्यकताक लेल विशिष्ट मॉडल प्रदान करैत अछि:

Flash v2.5 (फ्लैश वी2.5): ई अत्यन्त कम विलम्बता (Ultra-low latency, ~75ms) वला मॉडल अछि। ई 32 भाषाक समर्थन करैत अछि आ 40,000 वर्णक सीमा प्रदान करैत अछि। डेवलपरसँ जुड़ल काज जतय गति महत्वपूर्ण अछि, ततय एकर उपयोग होइत अछि।

Turbo v2.5 (टर्बो वी2.5): ई उच्च गुणवत्ता आ कम विलम्बता (~250ms-300ms) क उत्कृष्ट सन्तुलन प्रदान करैत अछि।

Eleven Multilingual v2: ई 29 भाषामे कथावाचन, ऑडियोबुक आ पोस्ट-प्रोडक्शन लेल सबसँ यथार्थवादी आ भावनात्मक रूपसँ समृद्ध मॉडल अछि। ई 10,000 वर्ण धरि लम्बा-प्रारूप जनरेशनक लेल उपयुक्त अछि।

चित्र-विवरण 6: इलेवनलैब्स (ElevenLabs) क वॉइस स्टूडियो अन्तरफलक। एक दिस 'Settings' मे 'Stability' आ 'Clarity' स्लाइडर देखा रहल अछि, आ बीचमे टेक्स्ट बॉक्समे देवनागरी लिपिमे किछु पाठ लिखल अछि।

स्वर-अनुकरण (Voice Cloning)

स्वर-अनुकरण तकनीकसँ कोनो व्यक्तिक स्वरक सटिक डिजिटल प्रतिकृति बनायल जा सकैत अछि। एकर उपयोग पॉडकास्ट, वीडियो वॉयसओवर वा ऑडियोबुक लेल कएल जाइत अछि।

तत्काल स्वर-अनुकरण (Instant Voice Cloning): मात्र 1 सँ 5 मिनट धरिक स्पष्ट ऑडियो नमूनाक आधार पर किछु सेकण्डमे स्वरक प्रतिकृति तैयार कएल जा सकैत अछि।

व्यावसायिक स्वर-अनुकरण (Professional Voice Cloning): एकर लेल लगभग 30 मिनटसँ अधिकक साफ ऑडियो नमूना आवश्यक होइत अछि। ई मॉडल स्वरक हर सूक्ष्म विवरण, पिचक उतार-चढ़ाव, आ भावनात्मक लय कें पकड़ैत अछि। ई तकनीक उद्यम-स्तरक सुरक्षा (Enterprise-grade security) द्वारा सुरक्षित अछि, जाहिमे एन्क्रिप्शन आ SOC 2, HIPAA जेहन अनुपालन शामिल अछि।

भावनात्मक नियन्त्रण आ सन्दर्भ

आधुनिक ए.आइ. स्वर केवल रोबोट जँका पाठ्य नहि पढ़ैत अछि, बरु पाठ्यमे निहित भावनात्मक संकेत (Emotional cues) पर प्रतिक्रिया दैत अछि। टेक्स्टमे दिशात्मक टैग (tags) जोड़लासँ डिलीवरीमे आश्चर्यजनक परिवर्तन आबि जाइत अछि। उदाहरण स्वरूप:

"स्वर एक क्षण लेल रुकि गेल, [softly] जेना किछु सोचि रहल हो।" (The voice paused for a moment, [softly] as if gathering its thoughts)।

[laughs warmly] (उष्णतासँ हँसैत) वा [sighs contentedly] (सन्तोषजनक साँस लैत) जेहन टैग जोड़लासँ ए.आइ. स्वरमे प्राकृतिक हँसी वा साँसक ध्वनि उत्पन्न होइत अछि, जे श्रोता कें एकटा मानव उपस्थितिक बोध कराबैत अछि।

भारतीय उद्यम लोकनि एहि तकनीकक उपयोग ग्राहक सेवा (टेलिकॉम, ई-कॉमर्स), स्वास्थ्य सेवामे अप्वाइंटमेंट शेड्यूलिंग, आ बहुभाषी ऑडियोबुक जनरेशनमे उच्च मात्रामे कए रहल छथि। इलेवनलैब्सक प्लिवो (Plivo), ओजोनेटेल (Ozonetel) सन भारतीय टेलीफोनी प्रदातासँ नेटिव इन्टीग्रेशन अछि।

डिजिटल अवतार आ लिप-सिंक (Lip-sync)

जखन वीडियो आ ऑडियो दुनू जनरेट भऽ जाइत अछि, तँ अगिला महत्वपूर्ण चरण दुनू कँ निर्बाध रूपसँ मिलायब (Syncing) अछि, विशेष रूपसँ जखन वीडियोमे कोनो डिजिटल अवतार वा पात्र बाजि रहल हो।

लिप-सिंक तकनीक आ अवतार

डिजिटल अवतार लेल सटीक लिप-सिंक प्राप्त करब परम आवश्यक अछि। ए.आइ. लिप-सिंक उपकरण जेना हेजेन (HeyGen), वेव2लिप (Wav2Lip) आ सिंक लैब्स (Sync Labs) कोनो भाषाक जनरेटेड ऑडियो कँ वीडियोक पात्रक होंठक गतिसँ सटीक रूपसँ मिला सकैत अछि। ई तकनीक वीडियो कँ बहुभाषी बनएबामे (Localisation) बहुत काज आवैत अछि।

सिनेमाटिक अवतार लेल कैमरा लेन्सक चुनाव सेहो महत्वपूर्ण अछि:

सामान्य बातचीत आ प्रतिक्रियाक लेल 35–50mm लेन्सक रूप (look) प्रयुक्त कएल जाइत अछि।

परिवेश देखाबेबाक लेल (Context shots जतय अधिक पृष्ठभूमि देखाबक अछि) 24–35mm लेन्स उपयुक्त होइत अछि।

ब्यूटी क्लोज-अप वा कोनो उत्पाद प्रदर्शन (Product handoffs) लेल 50–85mm लेन्सक निर्देश देल जाइत अछि।

फ्रेम रेट आ क्लिप अवधि

एकटा आम भ्रान्ति अछि जे उच्च फ्रेम रेट (जेना 60 fps) वीडियो कँ अधिक सिनेमाटिक वा 'स्मूथ' बनाबैत अछि। यद्यपि, वर्तमान टेक्स्ट-टू-वीडियो मॉडल मुख्य रूपसँ 24–30 fps वितरणक लग प्रशिक्षित अछि।

फ्रेम रेट (Frame Rate): सिनेमाटिक अनुभव आ प्राकृतिक मोशन ब्लर (motion blur) लेल 24 fps क उपयोग कएल जेबाक चाही। 60 fps क माँग करलासँ ए.आइ. इण्टरपोलेशन (interpolation) करय लागैत अछि, जकरासँ अस्वाभाविक मोशन ब्लर आ आर्टिफैक्ट्स (artifacts) आबि सकैत अछि। 25 वा 30 fps क उपयोग मात्र ओहि समय करबाक चाही जखन सोशल प्लेटफार्म वा ब्रॉडकास्टक तहिना आवश्यकता हो।

क्लिप अवधि (Clip Length): क्लिप कतेक लम्बा हो, ई प्लेटफार्म पर नहि, बरु ए.आइ. मॉडलक कालिक स्थिरता (temporal stability) पर निर्भर करैत अछि।

2-3 सेकण्डक क्लिप स्थिर होइत अछि मुदा भावनात्मक परिवर्तन देखाबय लेल बहुत छोट होइत अछि।

4-6 सेकण्डक क्लिप कथा स्पष्टता आ स्थिरताक बीच सर्वश्रेष्ठ सन्तुलन बनाबैत अछि।

8+ सेकण्डक क्लिपमे पात्रक पहचान (identity), प्रकाश व्यवस्था वा संयोजनमे अस्थिरता (drift) आबय लागैत अछि।

अतः, 20 सेकण्डक वीडियो बनेबा लेल 4-5 सेकण्डक चारि वा पाँच अलग-अलग क्लिप जनरेट कए ओकरा कोनो सम्पादन सॉफ्टवेयर (जेना CapCut) मे जोड़ब अधिक विश्वसनीय रणनीति मानल जाइत अछि।

नकारात्मक प्रॉम्प्टक अनुकूलन (Negative Prompts)

वीडियोमे अनावश्यक त्रुटि दूर करबा लेल एकटा छोट मुदा प्रभावी नकारात्मक प्रॉम्प्ट सूची (Negative list) बहुत काज आवैत अछि:

"no camera shake" (कोनो कैमरा कम्पन नहि)

"no jump cuts" (कोनो जम्प कट नहि)

"no glitching limbs or faces" (कोनो विकृत अंग वा चेहरा नहि)

"no flickering shadows" (छायामे कोनो अस्थिरता नहि, ई विशेषकर तखन उपयोगी अछि जखन प्रकाश अस्थिर लागय)

"no extreme facial expressions" (यदि अवतारक चेहरा बीच शॉटमे विकृत भऽ रहल हो तँ एकर प्रयोग करू)।

"no text overlays" (यदि अहाँ नहि चाहैत छी जे ए.आइ. स्वयं कोनो अस्पष्ट टेक्स्ट वीडियो पर लिखि दय)।

ए.आइ. वीडियो मूल्यांकन आ VBench-2.0

कोनो उत्पन्न वीडियो उत्कृष्ट अछि वा नहि, एकर वस्तुनिष्ठ मूल्यांकन कोना कएल जाए? एहि उद्देश्यसँ अनुसन्धानकर्ता लोकनि 'VBench' (वी-बेंच) नामक एकटा मूल्यांकन बेंचमार्क (Benchmark) विकसित केने छथि। जखन अहाँ वीडियो जनरेट करब, तँ अहाँकेँ ज्ञात हेबाक चाही जे तकनीकी दृष्टिमे एकटा 'खराब' वीडियो ककरा कहल जाइत अछि।

सतही विश्वसनीयता सँ आन्तरिक विश्वसनीयता धरि

प्रारम्भिक ए.आइ. मॉडल मात्र "सतही विश्वसनीयता" (Superficial faithfulness) पर काज करैत छल, अर्थात् वीडियो देखबामे नीक लागय आ टेक्स्ट प्रॉम्प्टसँ मेल खाए (जेना- फ्रेम सौन्दर्यशास्त्र, कालिक चिकनाई (temporal smoothness), आ साधारण टेक्स्टक पालन) । मुदा वास्तविकतामे ई मॉडल एहि बातसँ अनभिज्ञ छल जे भौतिक विज्ञानक नियम वा मानव शरीर रचना कोना काज करैत अछि।

एहि अन्तर कैं पाटबा लेल VBench-2.0 प्रस्तुत कएल गेल अछि, जे ए.आइ. वीडियो जेनरेटरक "आन्तरिक विश्वसनीयता" (Intrinsic faithfulness) क मूल्यांकन करैत अछि। आब वीडियो मात्र दृष्टिसँ नीक नहि लागय, बरु ओकरा वास्तविक दुनियाक भौतिक आ तार्किक नियमकेँ सेहो पालन करबाक चाही। VBench-2.0 मे पाँच मुख्य आयाम (Dimensions) अछि:

मानव निष्ठा (Human Fidelity): ई जनित वीडियोमे मानव आकृतिक संरचनात्मक शुद्धता (structural correctness) आ कालिक स्थिरता (temporal consistency) क मूल्यांकन करैत अछि।

संरचनात्मक शुद्धता: ई देखैत अछि जे कोनो पात्रक शारीरिक अंग जेना हाथ-पैर अचानक गायब वा विकृत तँ नहि भऽ रहल अछि। जेना- कोनो पात्र घूमि रहल अछि तँ ओकर शरीर अस्वाभाविक रूपसँ विकृत नहि हो (जकरा 'Thousand-hand yoga' प्रभाव कहल जाइत अछि) ।

कालिक स्थिरता (Identity & Clothes): चेहराक पहचानक लेल ArcFace आ RetinaFace क उपयोग कएल जाइत अछि जाहिसँ ज्ञात भऽ सकै जे पूरा वीडियोमे व्यक्ति ओही अछि वा बदलि गेल। कपड़ाक स्थिरता जांचय लेल LLaVA-video-7B मॉडलक प्रयोग होइत अछि, जे ई सुनिश्चित करैत अछि जे शॉट बदललापर पात्रक कपड़ाक रंग वा डिजाइन नहि बदलि जाए।

रचनात्मकता (Creativity): एकटा ही प्रॉम्प्टसँ मॉडल कतेक भिन्न-भिन्न (विविध) परिणाम दऽ सकैत अछि, एकर जांच पूर्व-प्रशिक्षित VGG-19 फीचर प्रतिनिधित्वसँ कएल जाइत अछि। ई विविधता (Diversity) कैं मापबैत अछि।

नियन्त्रणीयता (Controllability): ई मापैत अछि जे मॉडल कैमराक गति वा विशिष्ट निर्देश (जेना पैन, ज़ूम) कैं कतेक सटिकतासँ पालन करैत अछि।

भौतिकी (Physics): वीडियोमे गुरुत्वाकर्षण, प्रकाशक परावर्तन आ अन्य भौतिक नियमक पालन भऽ रहल अछि वा नहि।

सामान्य ज्ञान (Commonsense): ई सुनिश्चित करैत अछि जे वीडियो तार्किक रूपसँ सत्य प्रतीत हो।

चित्र-विवरण 7: VBench लीडरबोर्डक ग्राफिकल प्रतिनिधित्व, जतय विभिन्न ए.आइ. मॉडल (जेना Runway, Sora) कैं 0.3 सँ 0.8 क नॉर्मलाइज्ड स्केल पर मानव निष्ठा, रचनात्मकता आ भौतिकीक मापदण्ड पर रेट कएल गेल अछि।

एकर अतिरिक्त, VBench++ न केवल टेक्स्ट-टू-वीडियो (T2V) बरु इमेज-टू-वीडियो (I2V) मॉडलक लेल सेहो मूल्यांकन समर्थित करैत अछि आ ई पूर्ण रूपसँ ओपन-सोर्स अछि जाहिमे अडेप्टिव आस्पेक्ट रेशियो (adaptive aspect ratio) क सुविधा अछि। ई मॉडलक विश्वसनीयता (trustworthiness) कैं सेहो परखैत अछि।

सम्पूर्ण निर्माण प्रक्रियाक व्यावहारिक कार्यप्रवाह (Practical Workflow)

एकटा सफल वीडियो निर्माता बनबाक लेल केवल औजारक जानकारी पर्याप्त नहि अछि, बरु सम्पूर्ण कार्यप्रवाह कैं सुव्यवस्थित रूपसँ लागू करब आवश्यक अछि। ई 'स्वयं सीखू' अध्याय एहि सम्पूर्ण प्रक्रिया कैं एक ठाम प्रस्तुत करैत अछि।

एपीआई आ उन्नत मापदण्ड (API Integration)

तकनीकी उपयोगकर्ता किछु उन्नत वीडियो-औजारक एपीआई (API) द्वारा अनुपात, गुणवत्ता, आउटपुट-प्रारूप, कैमरा-व्यवहार आ क्लिप-श्रृंखला जेहन मापदण्ड नियन्त्रित कऽ सकैत छथि। अलग-अलग औजारमे नाम आ उपलब्धता बदलैत अछि; नीचाँक उदाहरणसभ अवधारणा बुझबाक लेल अछि:

aspect_ratio: 16:9 (यूट्यूब लेल), 9:16 (रील्स लेल), 1:1, 4:3, 21:9 (सिनेमाटिक) वा 'adaptive' जाहिसँ मॉडल पहिल फ्रेम वा प्रॉम्प्टक आधार पर स्वयं अनुपात तय करय।

bitrate_mode: 'standard' (CRF 18) वा 'high' (CRF 11) । हाई मोड 3-5 गुना पैघ फाइल बनबैत अछि मुदा एहिमे विवरण अत्यन्त स्पष्ट रहैत अछि।

output_format: .mov (yuv444p / yuv444p10le) रंगक उच्च निष्ठा (color fidelity) आ बार-बार क्लिप कैं विस्तार देबाक (extend/edit chains) लेल सर्वोत्तम अछि, जखन कि .mp4 (yuv420p) सोशल मीडिया पर सीधा अपलोड करबा लेल डिफ़ॉल्ट मानक अछि।

camera_fixed: एकरा 'true' करलासँ मॉडल लॉक-ऑफ कैमराक दिस बायस्ड (biased) भऽ जाइत अछि, अर्थात् कैमरा बिल्कुल नहि घूमत।

return_last_frame: ई 'true' केलासँ अन्तिम फ्रेम एकटा वाटरमार्क-मुक्त PNG क रूपमे प्राप्त होइत अछि, जकरा वापस इनपुटमे दऽ कैं ओतयसँ नव वीडियो (chaining) जनरेट कएल जा सकैत अछि, जाहिसँ 4 सेकण्डक क्लिप कैं 20 सेकण्ड धरि खींचल जा सकैत अछि।

अन्तिम संयोजन (Assembly): चरण-दर-चरण

मैथिलीमे अपन ए.आइ. वीडियो बनएबाक सम्पूर्ण प्रक्रिया एहि प्रकार अछि:

कहानी आ प्रॉम्प्टक विकास: सभसँ पहिने अपन विषय चुनू आ प्रॉम्प्ट इंजीनियरिंग (अध्याय 2) क सूत्रक आधार पर स्पष्ट अंग्रेजी प्रॉम्प्ट तैयार करू।

वीडियो जनन: उपलब्ध उपयुक्त पाठ-सँ-वीडियो वा चित्रसँ-वीडियो औजारक उपयोग कऽ छोट-छोट क्लिप बनाउ। यदि कोनो क्लिपमे विकृति, निरन्तरता-दोष वा अस्वाभाविक गति देखाइत अछि तँ गुणवत्ता-मापदण्ड ध्याने रखैत ओकरा अस्वीकार कऽ नव क्लिप बनाउ।

ऑडियो उत्पादन: अपन रिकॉर्ड कएल मैथिली स्वर वा उपयुक्त पाठ-सँ-स्वर औजारक प्रयोग करू। कृत्रिम स्वर उपयोग करैत छी तँ उच्चारण, नाम, विराम, भाव आ सहमति जाँचू; स्वर-अनुकरण केवल स्पष्ट अधिकार आ अनुमति रहने करू।

लिप-सिंक आ अवतार: यदि वीडियोमे बाजैत पात्र अछि तँ उपयुक्त लिप-सिंक वा अवतार-औजारक माध्यमसँ दृश्य आ ऑडियो कैं मिलाउ; ओठक गति, भाव, समय-साम्य आ पात्रक पहचान अलग-अलग जाँचू।

अन्तिम सम्पादन (Post-production): उपयुक्त टाइमलाइन-सम्पादकमे सभ छोट क्लिप, वाचन, ध्वनि-प्रभाव, पृष्ठभूमि-सङ्गीत, शीर्षक आ उपशीर्षक एक ठाम जोड़ू; अन्तिम निर्यातसँ पहिने दृश्य-निरन्तरता, ध्वनि-सन्तुलन आ पाठक शुद्धता जाँचू (Dancyger, २०२६; Clark आदि, २०१९)।

चित्र-विवरण 8: CapCut (डेस्कटॉप संस्करण) क टाइमलाइन, जतय वीडियोक 4 छोट-छोट क्लिप, ओकर नीचाँ ElevenLabs सँ उत्पन्न ऑडियो ट्रैक, आ अंतमे बैकग्राउंड म्यूजिक (BGM) लेयर सजल अछि।

निष्कर्ष

उपरि देल अध्ययनसँ स्पष्ट अछि जे ए.आइ.-आधारित वीडियो निर्माण आब व्यवहारिक रचनात्मक औजारक रूपमे स्थापित भऽ रहल अछि। जनरेटिव मॉडल दृश्य, गति, स्वर आ शैलीक अनेक रूप शीघ्र उत्पन्न कऽ सकैत अछि, मुदा भौतिक यथार्थ, पात्र-निरन्तरता, तथ्य, सांस्कृतिक सटीकता आ नैतिक निर्णयमे मानवीय जाँच आवश्यक रहैत अछि (Enochs, २०२६; Anderson, २०२६; Hutson आ Smith, २०२५)।

उचित प्रॉम्प्ट इंजीनियरिंगक सन्तुलन, कैमराक कोण आ गति पर नियन्त्रण, उपयुक्त फ्रेम रेट (24 fps), आ सटीक प्रकाश व्यवस्थाक ज्ञानक सङ्ग कोनो नव उपयोगकर्ता एहि 'स्वयं सीखू' पोथीक सहायतासँ मैथिली वा कोनो अन्य भाषामे उच्च-गुणवत्ता वला सिनेमाटिक वीडियोक निर्माण कए सकैत छथि। ई तकनीक निर्माता लोकनिकेँ असीमित स्वतन्त्रता आ नियन्त्रण प्रदान करैत अछि, जाहिसँ ओ अपन कल्पना केँ तकनीकी पटल पर पूर्ण रूपसँ जीवन्त रूप दऽ सकैत छथि।

चयनित ग्रन्थ-सूची

एहि ग्रन्थ-सूचीमे केवल पुस्तक-रूप स्रोत सम्मिलित अछि। जनरेटिव मीडिया, प्रॉम्प्ट-लेखन, दृश्य-रचना, सम्पादन, ध्वनि, वृत्तचित्र, उत्तर-उत्पादन आ ए.आइ.-नैतिकताक स्थायी सिद्धान्तक लेल ई ग्रन्थ उपयोगी अछि। औजार-विशेषक बटन, योजना, मूल्य आ संस्करण क्षणभंगुर होइत अछि; तेँ पोथीक मुख्य जोर स्थायी कौशल आ कार्यप्रवाह पर राखल गेल अछि।

Anderson, Iain. 2026. *AI for Creative Production: A Handbook for the Ethical Use of AI in Creating and Processing Text, Images, Video, and Audio*. 1st ed. Packt Publishing.

Block, Bruce. 2021. *The Visual Story: Creating the Visual Structure of Film, TV, and Digital Media*. 3rd ed. Routledge.

Bowen, Christopher. 2024. *Grammar of the Shot*. 5th ed. Routledge.

Brown, Blain. 2022. *Cinematography: Theory and Practice: For Cinematographers and Directors*. 4th ed. Routledge.

Chattopadhyay, Budhaditya. 2024. *Sound in Indian Film and Audiovisual Media: History, Practices and Production*. 1st ed. Routledge.

Clark, Barbara, Susan Spohr, Dawn Higginbotham, and Kumari Bakhru. 2019. *The Guide to Managing Postproduction for Film, TV, and Digital Distribution: Managing the Process*. 3rd ed. Routledge.

Coeckelbergh, Mark. 2020. *AI Ethics*. The MIT Press.

Dancyger, Ken. 2026. *The Technique of Film and Video Editing: History, Theory, and Practice*. 7th ed. Routledge.

Dubber, Markus D., Frank Pasquale, and Sunit Das, eds. 2020. *The Oxford Handbook of Ethics of AI*. Oxford University Press.

Enochs, Joseph. 2026. *Video Generation with AI*. O’Reilly Media.

Farnell, Andy. 2010. *Designing Sound*. The MIT Press.

Foster, David. 2023. *Generative Deep Learning*. 2nd ed. O’Reilly Media.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. The MIT Press.

Hutson, James, and Andrew Smith. 2025. *Cinematic Algorithms: The Rise of Generative AI in Video Art and Visual Culture*. CRC Press.

Khan, Hussin. 2024. *Virtual Filmmaking with Unreal Engine 5: A Step-by-Step Guide to Creating a Complete Animated Short Film*. 1st ed. Packt Publishing.

Liao, S. Matthew, ed. 2020. *Ethics of Artificial Intelligence*. Oxford University Press.

Phoenix, James, and Mike Taylor. 2024. *Prompt Engineering for Generative AI*. 1st ed. O’Reilly Media.

Rabiger, Michael, and Courtney Hermann. 2026. *Directing the Documentary*. 8th ed. Routledge.

Rose, Jay. 2015. *Producing Great Sound for Film and Video: Expert Tips from Preproduction to Final Mix*. 4th ed. Routledge.

Russell, Stuart, and Peter Norvig. 2020. *Artificial Intelligence: A Modern Approach*. 4th ed. Pearson.

Sanseviero, Omar, Pedro Cuenca, Apolinário Passos, and Jonathan Whitaker. 2024. *Hands-On Generative AI with Transformers and Diffusion Models*. 1st ed. O’Reilly Media.

ई ग्रन्थ-सूची विषयगत पठनक लेल अछि; पोथीमे वर्णित कोनो एक व्यावसायिक औजारक अस्थायी संस्करण, मूल्य वा अन्तरफलक केँ स्थायी मानक मानबाक उद्देश्य नहि अछि।

विदेह

गजेन्द्र ठाकुर
सम्पादक — विदेह मैथिली ई-पत्रिका
www.videha.co.in
ISSN 2229-547X

Teach Yourself Series

Make Videos with AI

A Complete Practical Course from Script to Publication

With illustrations, screenshots, exercises, projects, technical expansion, and language refinement

Gajendra Thakur

Editor — Videha Maithili e-Journal

www.videha.co.in | ISSN 2229-547X

First Edition — 2026

Publication Details

© 2026 Gajendra Thakur. All rights reserved. No substantial portion of this book may be commercially reprinted, sold, or adapted without the author's written permission. For fair quotation in teaching, review, and research, the source must be clearly acknowledged.

The names, logos, and interfaces of the tools used in this book are trademarks of their respective companies. Screenshots are included in a limited manner solely to explain steps, support review, and teach digital literacy. Interfaces, plans, prices, and versions may change over time; therefore, instead of memorizing button names, this book emphasizes durable working principles, creative decision-making, ethics, and production workflow.

Neither the author nor the publisher guarantees that any subscription price, credit allowance, availability, or regional feature will remain unchanged. Before publication, check the current terms of service, privacy policy, copyright rules, consent requirements, and platform rules yourself.

Book format: A4 illustrated self-study guide. Language: English. First edition: 2026.

ISBN: 978-93-5943-562-6

Preface

In the past, making a video was thought to require a camera, lighting, a studio, actors, an editor, and a large budget. Artificial intelligence has not abolished that system, but it has certainly placed new tools in the hands of small creators. Today, with a clear idea, a well-structured script, a few images, carefully considered prompts, and basic editing skills, one can create educational videos, book introductions, folktales, promotional films, short films, song sequences, supporting visuals for documentaries, and short-form content for social platforms.

The purpose of this book is not to make you memorize the buttons of any one company, but to develop fundamental skills that remain useful even when the tools change. For this reason, each lesson presents the principle first, followed by screenshots, hands-on practice, self-checks, and a short project.

Maithili creators face two particular challenges: first, producing natural scripts and voices in the mother tongue; and second, avoiding the reduction of local culture to stereotypes or artificial decoration. This book trains you to use the environment, clothing, architecture, folk art, pronunciation, and social sensitivities of Mithila responsibly.

AI is an assistant, not the author. Final meaning, factual accuracy, ethical responsibility, and aesthetic judgment remain in human hands. The clearer your decisions, the more original the result will be.

How to Read This Book

The book is organized at three levels. At the first level, you will learn to make short videos on a mobile device or in a browser without prior experience. At the second level, you will gain greater control over storyboards, camera language, prompts, voice, and editing. At the third level, you will develop longer projects, ethical publishing practices, and a regular production system.

64. Read the objectives given at the beginning of each chapter.
65. Match the red numbers marked in the screenshots with the corresponding steps in the text.

66. Save the output of each exercise in a separate folder.
67. Do not treat the first prompt result as final; make three or four versions with small changes.
68. Complete at least one of the three projects in Chapter 15.
69. Complete the final 21-day programme and build your own portfolio of samples.

Organization of the Book

- Part 1 — Core Teach Yourself Course: from basic concepts to publication, projects, and the 21-day programme
- Part 2 — Expanded Explanation and Application: planning, storyboarding, tools, voice, avatars, music, Maithili content, and exercises
- Part 3 — Advanced Technical Study: leading AI video generators of 2026, prompt design, camera and lighting, voice cloning, lip-sync, VBench, and API workflows
- Appendices — prompt collection, worksheets, glossary, and tool guide; followed by a selected bibliography

Part 1
Core Teach Yourself Course

What AI Video Is — and What It Is Not

After this chapter, you will be able to —

• distinguish generative video from ordinary editing

• identify the strengths and limitations of AI

• choose an appropriate project

Key terms: generative video · text-to-video · image-to-video · human editing

AI video is moving-image content in which some or all of the visuals, motion, voice, subtitles, editing, or transformation have been produced with the help of machine-learning-based tools. This does not mean that an entire film is completed simply by writing one sentence. In practice, a successful video is the result of many small decisions.

Four Main Types

Text-to-video: you provide a written description of the scene; the tool creates a short clip from it.

Image-to-video: you provide a starting image; the prompt mainly describes motion, camera behavior, and atmosphere.

Video-to-video: the style, background, clothing, lighting, or motion of an existing clip is changed.

Assistive AI: used for scriptwriting, automatic subtitles, noise reduction, voice generation, music, color correction, or short-form extraction.

What AI Does Well

- It creates imaginative alternatives to B-roll that is unavailable or expensive to obtain.
- It adds movement, camera motion, and atmosphere to still images.
- It quickly produces several visual interpretations of the same idea.
- It speeds up the creation of scripts, titles, shot lists, and editing plans.
- It reduces the labour involved in voice generation, dubbing, subtitles, and resizing.

Where AI Is Weak

- In long scenes, continuity of faces, hands, objects, and clothing may break down.
- Written text, logos, maps, calculations, or historical facts may be wrong.
- Too many characters or several simultaneous actions in a single scene can cause confusion.
- Maithili pronunciation, local vocabulary, and cultural nuance cannot be assumed to be automatically correct.
- A scene that looks realistic is not necessarily true.

Core rule: use AI to create visuals, but verify facts yourself; use AI to create a voice, but listen to the pronunciation yourself; use AI to draft a script, but keep the meaning and style your own.

Self-Check

70. What is the basic difference between text-to-video and image-to-video?
71. Why is a long narrative video better created as a series of short clips rather than generated all at once?
72. Who is responsible for checking whether a realistic-looking AI scene is factually true?

Answer hints: 1. In text-to-video, both scene and motion are described; in image-to-video, the image determines the scene and the prompt mainly describes motion. | 2. Continuity, control, and regeneration are better. | 3. The human creator and publisher.

Workflow: From Idea to Publication

After this chapter, you will be able to —

• adopt a ten-step production system

• avoid losses caused by doing tasks in the wrong order

• preserve the output of every stage

Key terms: workflow · version · shot list · quality check

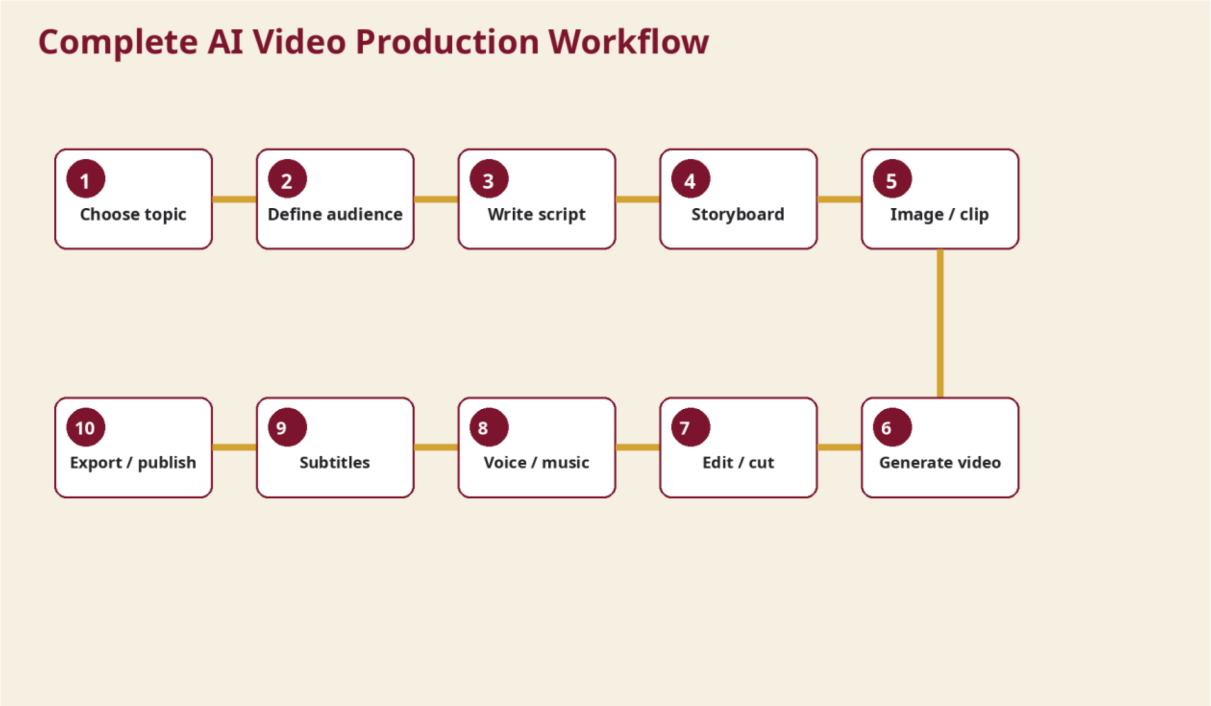


Figure 2.1 — Ten Steps in AI Video Production

Original teaching illustration

Why Sequence Matters

Beginners often open a tool and start generating scenes before doing anything else. Then the topic changes, the aspect ratio is wrong, the voice duration does not match the visuals, and many clips are wasted in the end. A proper sequence saves time, credits, and mental effort.

Topic: write in one sentence who the video is for and what it will say.

Audience: determine age, language, platform, and prior knowledge.

Script: write the voice text first, then add the visuals.

Storyboard: make a visual plan for each 3–8 second scene.

Assets: organize your own photos, logos, maps, audio, and sources.

Generation: first create a low-quality or short-duration test.

Editing: select the best clips and improve sequence, pacing, and transitions.

Voice/audio: keep speech clear, music lower, and effects restrained.

Subtitles: check Maithili spelling yourself; do not rely on automatically generated text.

Export/publication: choose the correct aspect ratio, resolution, name, description, and AI disclosure.

A Simple Rule for Version Names

Saving the same file repeatedly as "final" creates confusion. Add the date, project, and version: mithila_pokhari_v01_2026-08-04.mp4. After revisions, use v02, v03, and so on. You may add FINAL to an approved file, but do not delete the older versions.

Exercise 2.1 — One-Minute Video Plan

73. Choose a topic: "Makhana: From Pond to Plate," or a topic of your own interest.
74. Write the audience and purpose in one sentence.
75. Write the sequence of six scenes.
76. For each scene, write one line giving duration, image, and voice.

Exercise output: a one-page production plan that will become the basis for the script in the next chapter.

Tools, Accounts, and File Organization

After this chapter, you will be able to —

• prepare the minimum setup on mobile and computer

• choose the correct aspect ratio and resolution

• create a safe folder structure for assets

Key terms: browser · credits · aspect ratio · resolution · backup

Minimum Equipment

- A stable internet connection; generative video files are large.
- A modern browser; many web tools recommend a recent version of Chrome or Safari.
- At least 5–10 GB of free storage; even unsuccessful tests may be worth keeping.
- Headphones; small audio defects may not be audible on a phone speaker.
- A quiet room and an ordinary microphone for recording your own voice.

Folder Structure

Inside the main folder for each project, create folders named 01_script, 02_storyboard, 03_images, 04_generated_clips, 05_voice, 06_music_sfx, 07_edit, 08_export, and 09_sources. Numbered names keep the sequence clear at all times.

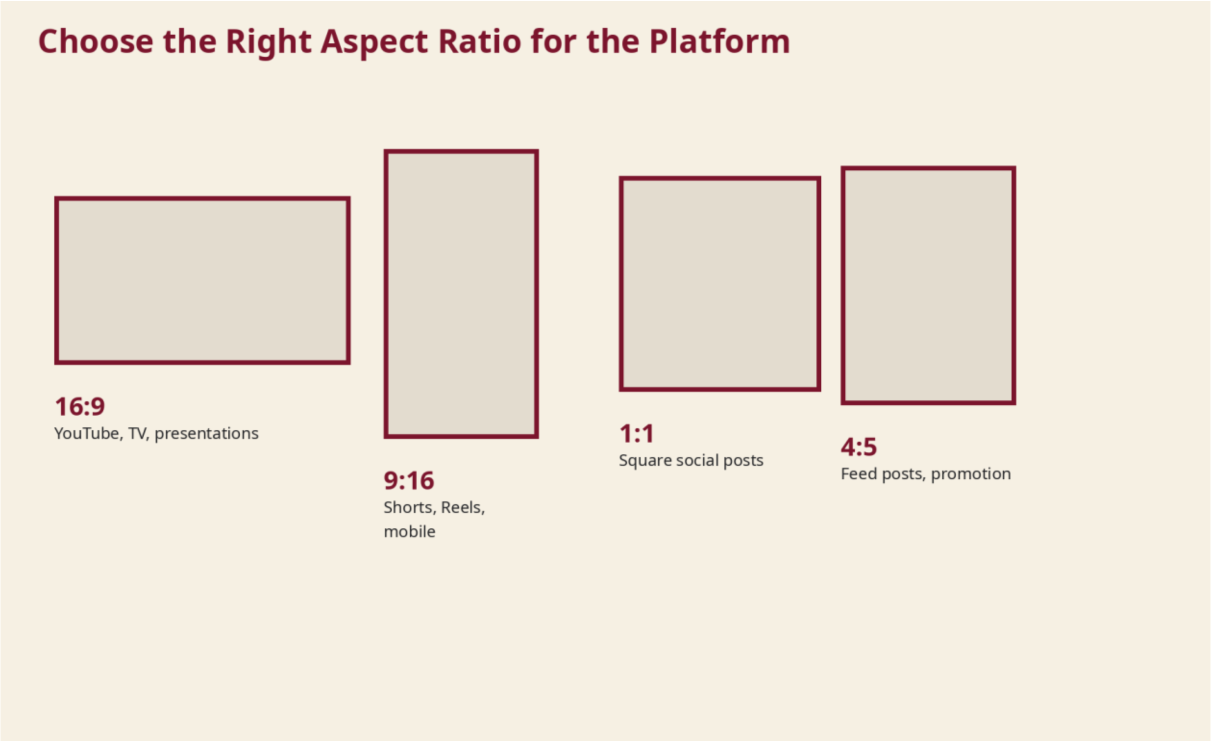


Figure 3.1 — Common Aspect Ratios by Platform

Original teaching illustration

Choose the Aspect Ratio First

A 16:9 horizontal video is suitable for YouTube, television, and lectures. A 9:16 vertical video is for mobile Shorts and Reels. Use 1:1 for square posts; 4:5 occupies more screen space in social feeds. Cropping a scene made for one ratio into another can push the main subject or text outside the frame.

Ways to Save Credits

- Begin with a 5-second test.
- Change only one thing at a time; do not alter the prompt, duration, and camera simultaneously.
- 720p is sufficient at the beginning; generate only the finally approved scene at higher quality.
- Before regenerating, name the defect: face, motion, lighting, background, or camera.
- Stop once you have three useful versions of the same scene; endless options are not creativity but indecision.

Safety: before uploading confidential manuscripts, unpublished personal photographs, identity documents, phone numbers, or private videos, read the tool’s privacy policy.

Ideas, Research, and Scriptwriting

After this chapter, you will be able to —

• turn a one-sentence idea into a video script

• write a hook, main section, and ending

• separate visuals and narration into different columns

Key terms: hook · narration · B-roll · scene duration · source

Begin with One Sentence

A successful idea contains three things: subject, audience, and outcome. Example: “Explain the basic idea of Mithila’s Panji genealogical tradition to school students in two minutes using simple visuals.” This sentence makes clear that the video is an introductory explanation, not a complete historical study.

Three-Part Structure

Opening hook: in the first 3–8 seconds, use a question, surprise, visual, or promise.

Main section: present one idea at a time, with a suitable visual for each idea.

Ending: provide a summary, a useful next step, a source, or a call to action.

Sample — 60-Second Script

Time	Voice-over	Visual	Text/Sound
0–5 seconds	“One seed, one pond, and the economy of an entire village — this is makhana.”	Aerial view of a pond at dawn	Title appears slowly
5–18	Makhana is cultivated in water...	Close-ups of farmer, leaves, flowers, and water	Soft sounds of water and birds
18–35	The fruit is collected and dried...	Sequence of hand-work processes	Three short labels
35–50	After roasting, the seed pops open into white makhana.	Roasting and popped makhana	Light rhythmic beat
50–60	Makhana is not merely food; it is part of Mithila’s water culture.	Plate, family, pond	Sources and closing

Rules for Writing a Script

- A sentence that is difficult to speak is also difficult to hear; write short sentences.

- If the image already makes something clear, do not repeat the same information in the voice-over.
- Make a separate list of numbers, names, and pronunciations.
- Note the source of every fact; this makes it easier to provide sources in the final description.
- For Maithili narration, read the punctuation, vowel length, and difficult conjunct letters aloud and check them yourself.

Exercise 4.1 — Your Own Script

77. Use the topic selected in Chapter 2.
78. Write a 60–90 second narration in 130–180 words.
79. Write the corresponding visual beside each sentence.
80. Read it aloud on a mobile phone and time it; if it exceeds the target duration, reduce the word count.

Exercise output: a timed, spoken-aloud, visually planned script.

Storyboard and Camera Language

After this chapter, you will be able to —

• divide a script into short scenes

• choose shot sizes and camera movements

• maintain continuity

Key terms: storyboard · close-up · wide shot · pan · dolly · continuity

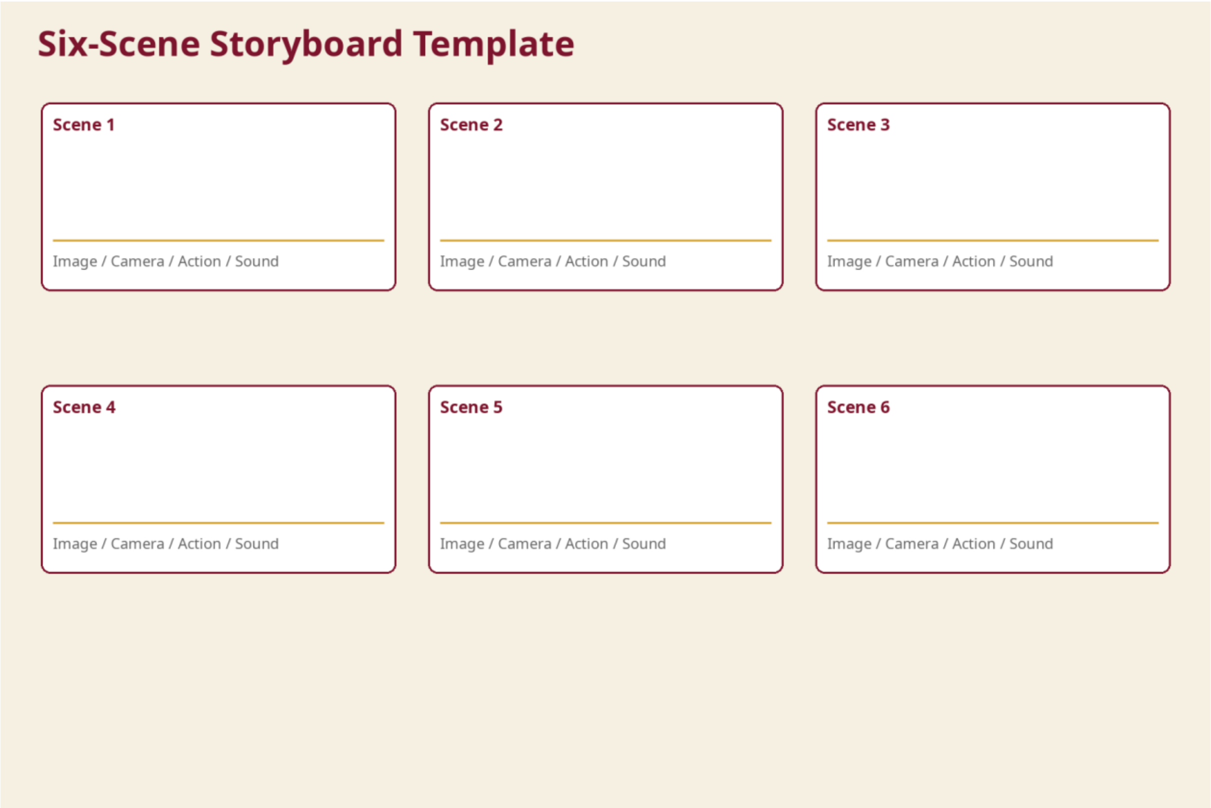


Figure 5.1 — Six-Scene Storyboard

Original worksheet; it may be printed and used

One Scene — One Main Action

Keep only one main action in a short generative clip. An instruction such as “the girl runs, opens a book, birds fly, and the camera circles around” will usually produce unstable results. Divide it into four scenes.

Common Shot Sizes

Extreme close-up: eyes, hands, letters, textures; for emotion or detail.

Close-up: face or main object; focuses the viewer’s attention.

Medium shot: character from the waist up; useful for dialogue and action.

Wide shot: full character and environment; introduces the location.

Extreme wide shot: landscape, village, building, or a large event.

Camera Movement

Pan: the camera turns left or right.

Tilt: the camera turns up or down.

Dolly in: the camera moves closer to the subject.

Dolly out: the camera moves backward away from the subject.

Tracking: the camera moves along with the subject.

Locked/static: the camera does not move; useful for education, products, and dialogue.

Continuity Sheet

If one character appears in several scenes, write the character’s age, face, hair, clothing, jewellery, expression, lighting, and location in a “character sheet.” Whenever possible, use the same reference image. Do not add unnecessary new adjectives to the prompt for every scene.

> Visual restraint: cinematic quality does not come from moving the camera all the time. Many powerful scenes are created with a locked camera, clear composition, and subtle human movement.

The Art of Writing Prompts

> After this chapter, you will be able to —
>
> • build seven-element prompts
>
> • write different instructions for text-to-video and image-to-video
>
> • refine results through repeated iteration
>
> Key terms: prompt · negative instruction · camera · style · seed

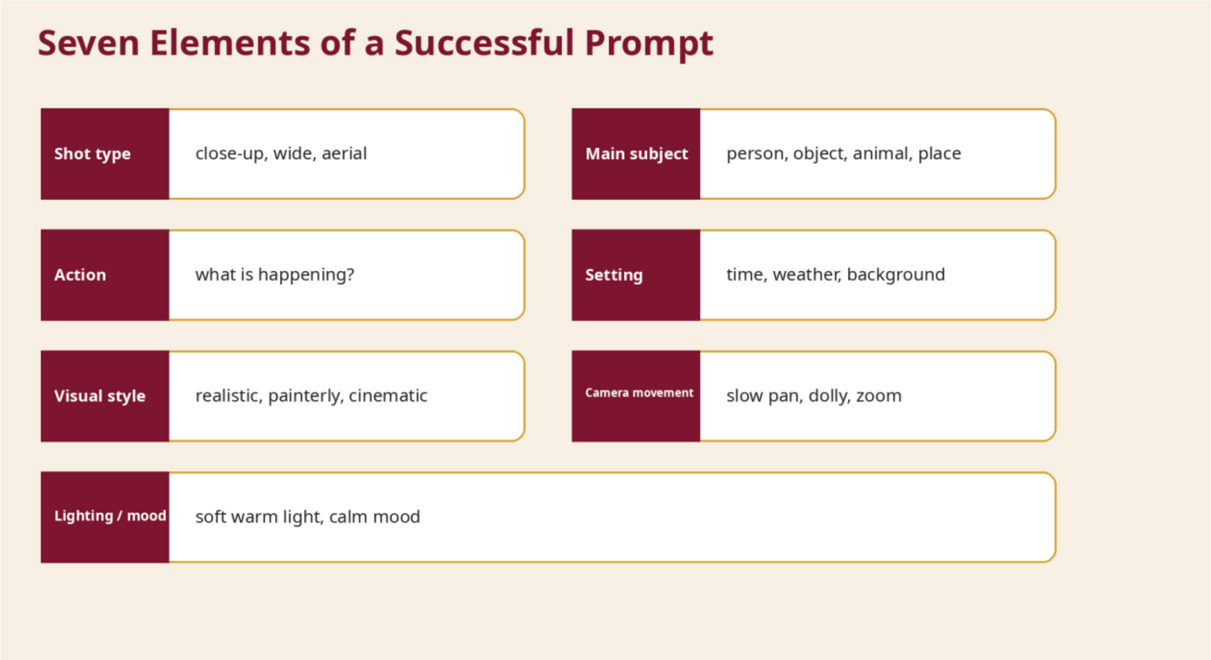


Figure 6.1 — Seven Elements of a Successful Prompt

Original teaching illustration

The Basic Prompt Formula

Shot type + main subject + clear action + place/time + visual style + camera movement + lighting/mood. The purpose of this sequence is not to decorate the language, but to make the decisions clear.

Sample 1 — Text-to-Video

> Prompt to enter in the tool: A wide cinematic dawn shot of a Mithila village pond, lotus leaves and makhana plants moving gently in the breeze, two fishermen in the far background, slow forward dolly, soft golden light, natural colors, realistic documentary style, calm water reflections, no text, no logo.

Maithili meaning: a wide cinematic dawn scene; a pond in a Mithila village, lotus leaves and makhana leaves moving gently in the breeze, two fishermen in the distance, the camera moving slowly forward, soft golden light, natural colours, realistic documentary-like style, calm reflections; no text or logo.

Sample 2 — Image-to-Video

> Motion-focused prompt: The woman remains seated and looks toward the pond. Her sari edge moves gently in the breeze. Small ripples spread across the water. The camera makes a very slow push-in. Keep facial features, clothing, background architecture and colors unchanged.

The image has already determined the subject, clothing, and location; therefore, instead of describing the face again, the prompt specifies the motion and what must not change.

Why Weak Prompts Fail

“Make a beautiful video”: “beautiful” is vague; there is no subject, action, or camera instruction.

Too many adjectives: “amazing, divine, magnificent, magical...” do not make visual decisions.

Too many actions: the sequence breaks down within a short duration.

Contradictory instructions: “locked camera” and “rapidly rotating camera” at the same time.

Written text: letters may become distorted in a generative scene; add text later in the editor.

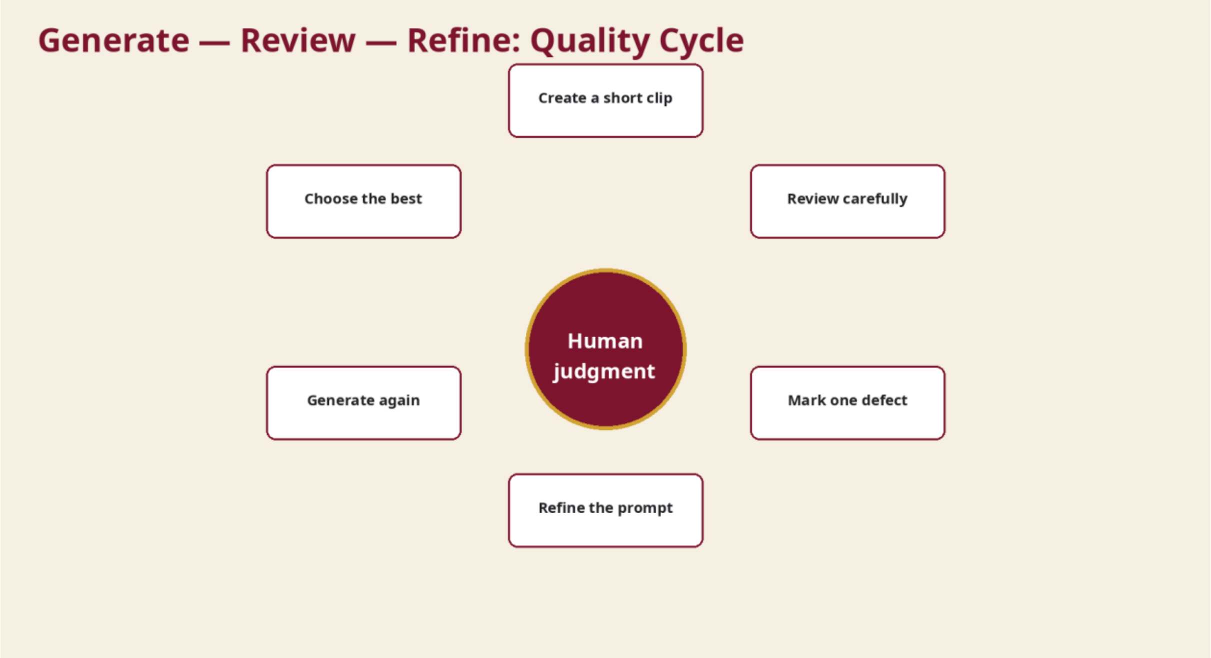


Figure 6.2 — Generate, Review, and Refine Cycle

Original teaching illustration

Change One Thing at a Time

If the camera is too fast in the first clip, add only “very slow, almost static camera.” If face, style, lighting, and duration are all changed together, it becomes difficult to know what caused the improvement. Save the prompt from every test.

Exercise 6.1 — Three Versions

81. Write a realistic documentary-style prompt for one scene.
82. Write a 2D illustration version of the same scene.
83. Write only an image-to-video motion prompt for the same scene.
84. Underline the words that changed in all three versions.

Exercise output: three controlled prompts for one subject that clearly show the difference between style and input type.

Text-to-Video: Web Tools

After this chapter, you will be able to —

• generate clips in web tools such as Runway Gen-4.5

• choose model, aspect ratio, duration, and motion
• avoid relying on outdated tutorials
Key terms: Runway · Firefly · model · duration · FPS

The names and appearance of web-based tools change rapidly, but the basic process remains almost the same: choose a model, write the prompt, set aspect ratio and duration, generate a test, identify defects, and then refine.

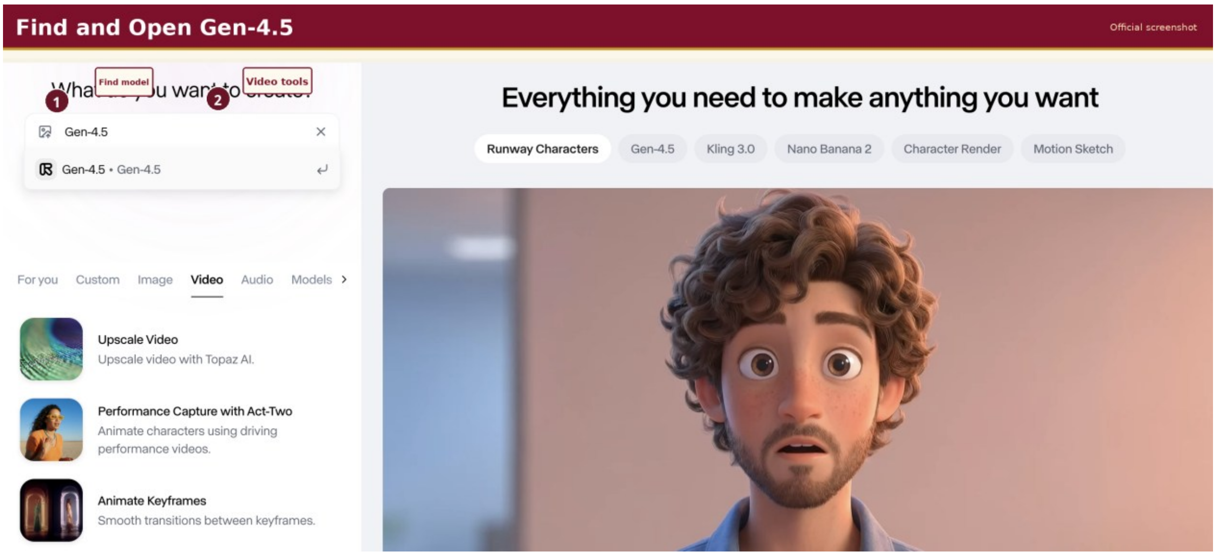


Figure 7.1 — Finding Gen-4.5 in Runway

Modern generative video tools such as Runway Gen-4.5 provide examples of both text-to-video and image-to-video workflows. In text-to-video, describe both the scene and the motion clearly; in image-to-video, concentrate mainly on motion, camera behavior, and continuity. Options for duration, resolution, and aspect ratio may vary by version.

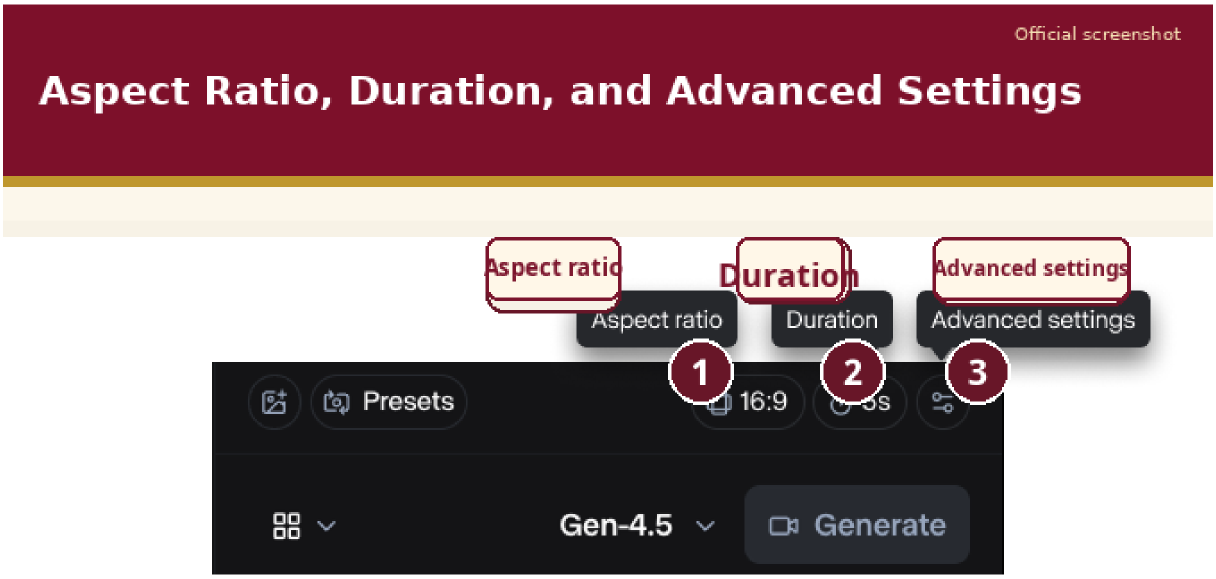


Figure 7.2 — Aspect Ratio, Duration, and Advanced Options

Step by Step

85. Open a new project and select Video or the relevant generative tool.

86. Choose the current video model from the model list.
87. For text-to-video, write a seven-element prompt.
88. Choose the aspect ratio according to the target platform.
89. Keep the first test to 5–8 seconds.
90. Generate it; watch the entire result, not just the first frame.
91. Download the best version and save it together with the prompt.

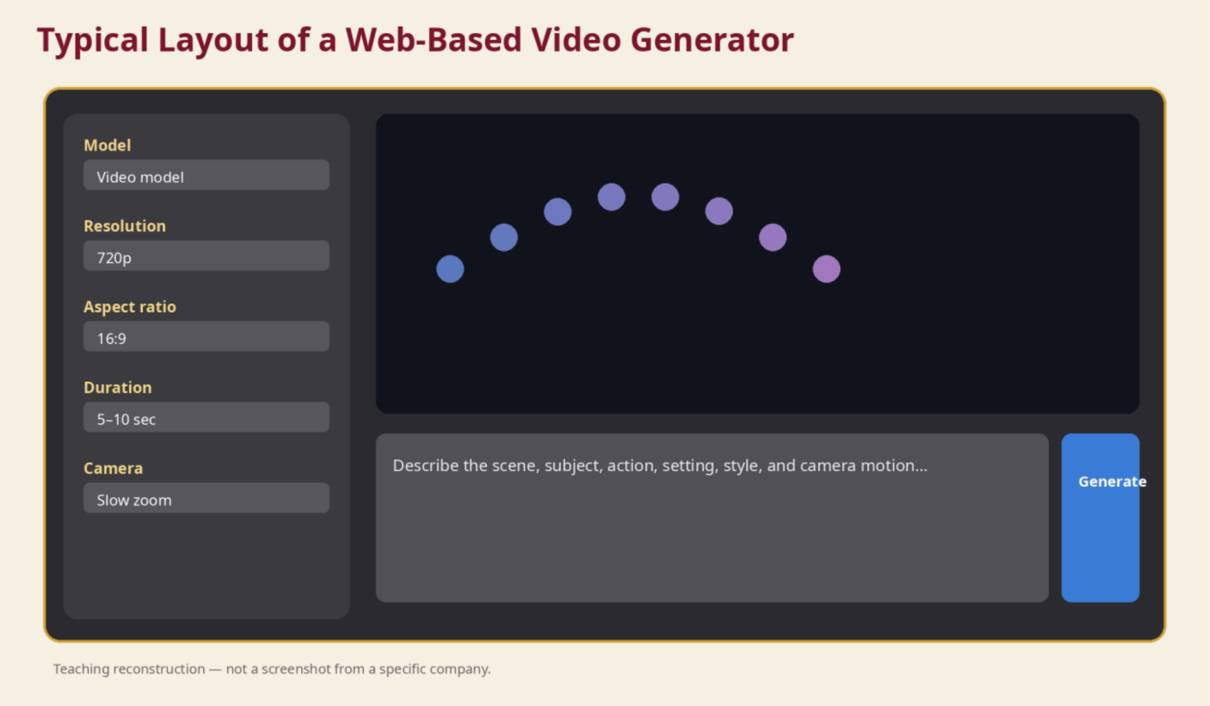


Figure 7.3 — Typical Layout of a Web Video Generator

Reconstructed example for explanation; not a screenshot from an actual company

Typical Adobe Firefly Workflow

Modern generative video tools generally let you choose a model, resolution, aspect ratio, reference image or reference video, camera options, and a prompt. In an effective prompt, it is useful to present shot type, subject or character, action, location, lighting, and visual style in a clear sequence (Phoenix and Taylor, 2024; Bowen, 2024).

Outdated Sora tutorial: the source text states that OpenAI’s Sora web and app experience ended on 26 April 2026 and that the Sora Videos API was scheduled to end on 24 September 2026 [9][10]. Therefore, do not spend time or money on old lessons built around obsolete “Sora app” buttons.

Self-Check

92. What two main things must be described in a text-to-video prompt?
93. Why should the first test be made at short duration and medium resolution?
94. What should you do first when you encounter an old screenshot or tutorial?

Answer hints: 1. Scene and motion. | 2. To save credits and refine quickly. | 3. Check the latest official help material.

Image-to-Video: Motion and Continuity

After this chapter, you will be able to —

• choose an appropriate starting image

• write motion-focused prompts

• keep faces and backgrounds stable

Key terms: reference image · first frame · motion · continuity · image rights

For beginners, image-to-video is often more controllable than text-to-video because the character, clothing, colour, and composition are already fixed in the source image. However, defects in a poor image can become more noticeable in the video.

Qualities of a Good Source Image

- The main subject should be clear and sufficiently large.
- Hands, face, and important objects should not be cropped.
- Lighting should come from a consistent direction and look natural.
- The background should not be excessively cluttered.
- The image aspect ratio should match the final video.
- The image should be your own, in the public domain, or used with written permission.

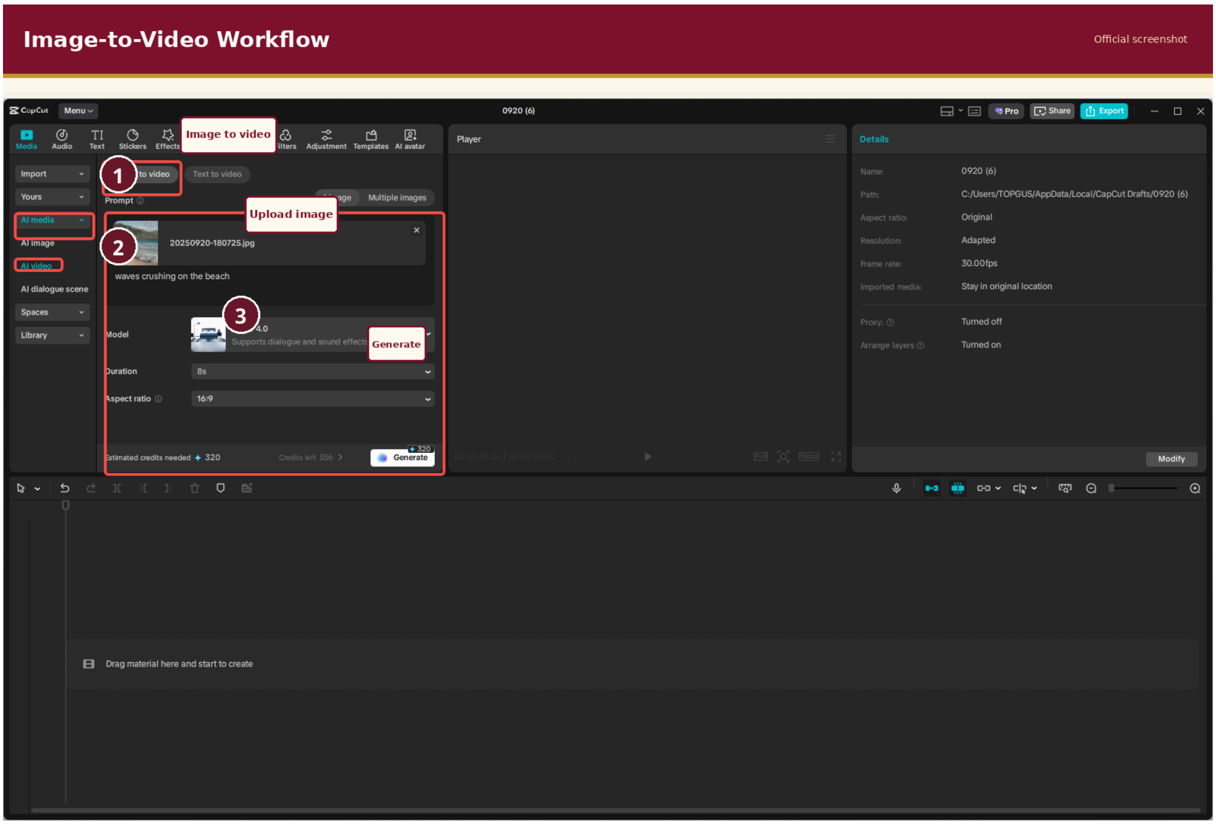


Figure 8.1 — Image-to-Video Option in CapCut

Four Levels of Motion Description

Subject motion: blinking, turning the head, walking, raising a hand.

Environmental motion: wind, water, smoke, leaves, clouds.

Camera motion: static, slow push-in, gentle pan.

Stability instruction: do not change the face, clothing, building, colours, or main objects.

Animating Mithila Painting

Instead of turning a Madhubani painting or line drawing into a photorealistic moving image, preserve its original flat composition and give it only subtle motion: let a fish's tail move gently, make water shimmer on a lotus leaf, keep the boundary lines fixed, and let the camera move forward extremely slowly. An instruction such as "keep the original flat painted style and linework unchanged" can help.

Exercise 8.1 — One Image, Three Motions

95. Choose one clean image that you have the right to use.
96. In the first version, describe only subtle movement in the wind or water.
97. In the second version, add one action by the subject.
98. In the third version, add slow camera movement.
99. In all three, check the stability of the face, clothing, and background.

Exercise output: three five-second clips and a comparison sheet.

Instant Full Video: CapCut Workflow

After this chapter, you will be able to —

• generate an initial video automatically from a script

• improve visuals, text, and music

• export with appropriate settings

Key terms: CapCut · AI Video Maker · script · timeline · export

Editors such as CapCut combine AI generation and conventional editing in one place. This allows a beginner to create an initial complete video from a script and then replace or refine each scene. In practice, it is useful to follow a sequence of pre-production, asset creation, timeline editing, audio, subtitles, and export (Clark et al., 2019; Dancyger, 2026).

Open the AI Video Maker Official screenshot

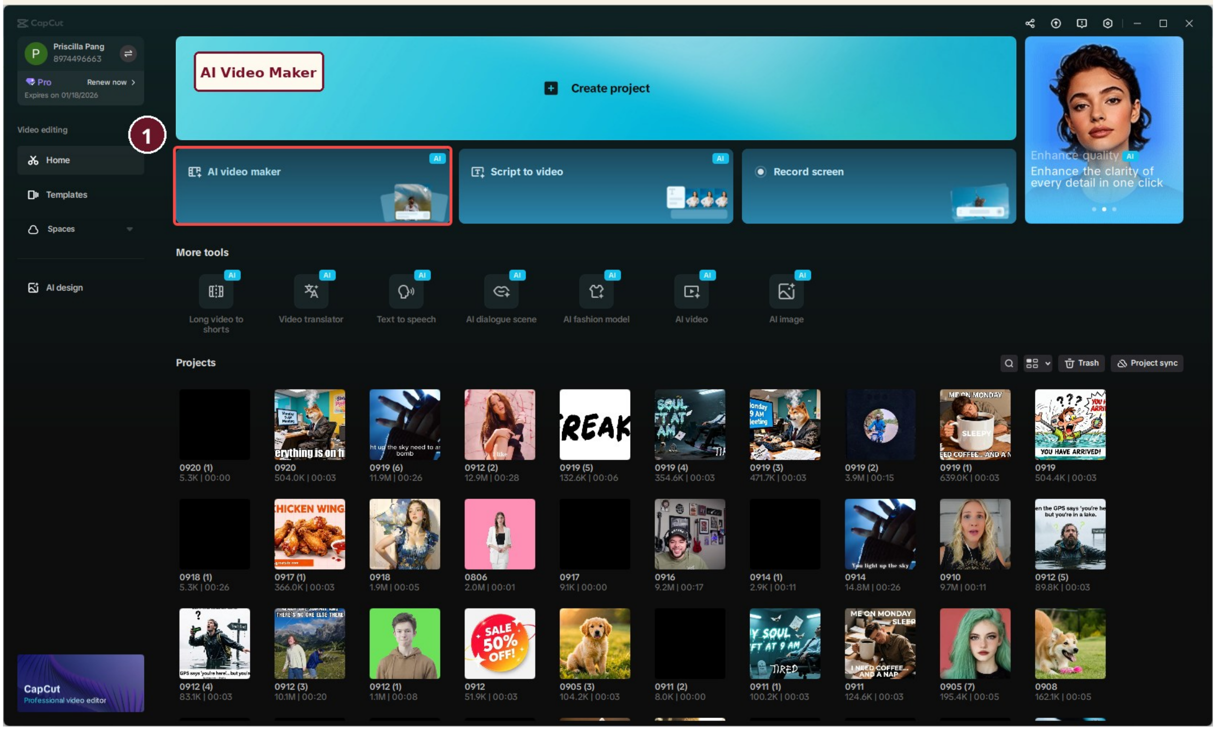


Figure 9.1 — Open the AI Video Maker

Enter Topic, Style, and Script Official screenshot

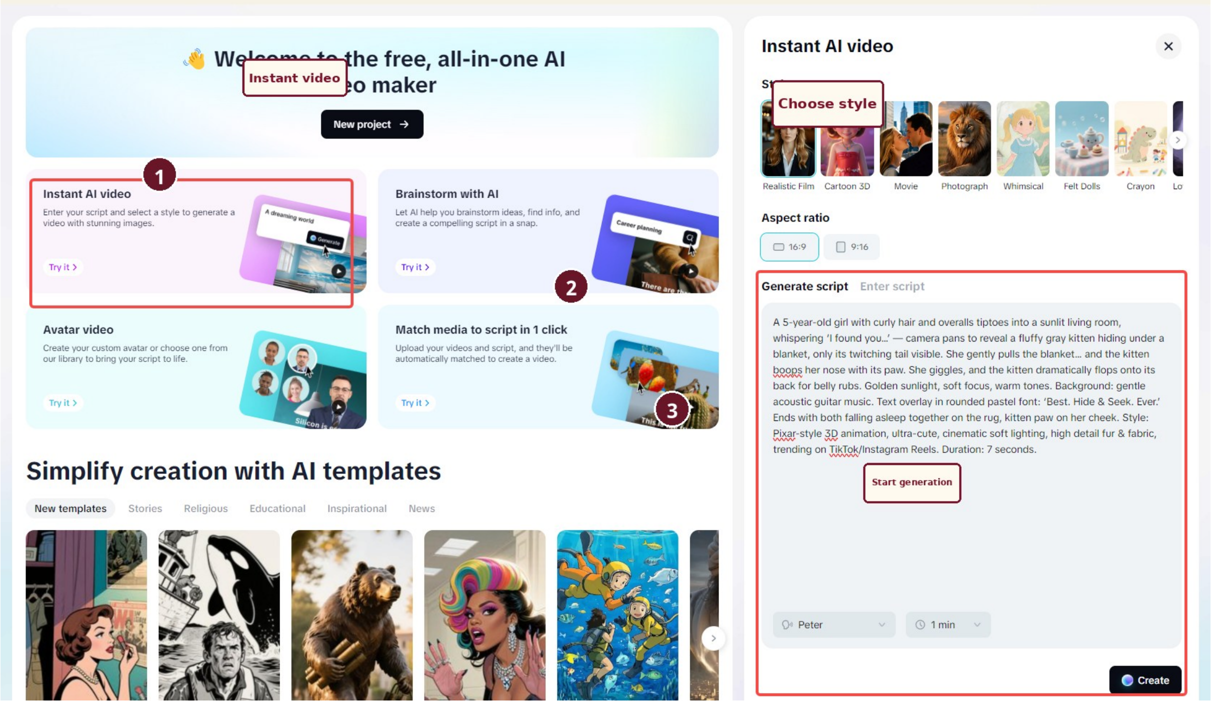


Figure 9.2 — Choose Style and Aspect Ratio, Then Enter the Script

Sequence for Making an Instant Video

100. Open the AI Video Maker.
101. Choose Instant AI Video or a similar option.

102. Choose a style; do not select a fashionable “trend” style that does not suit the project’s subject.
103. Choose a 16:9 or 9:16 aspect ratio.
104. Paste the script; check the scene divisions yourself before generating.
105. Generate the first version of the video.
106. Improve the image, duration, text, music, and order of every scene.

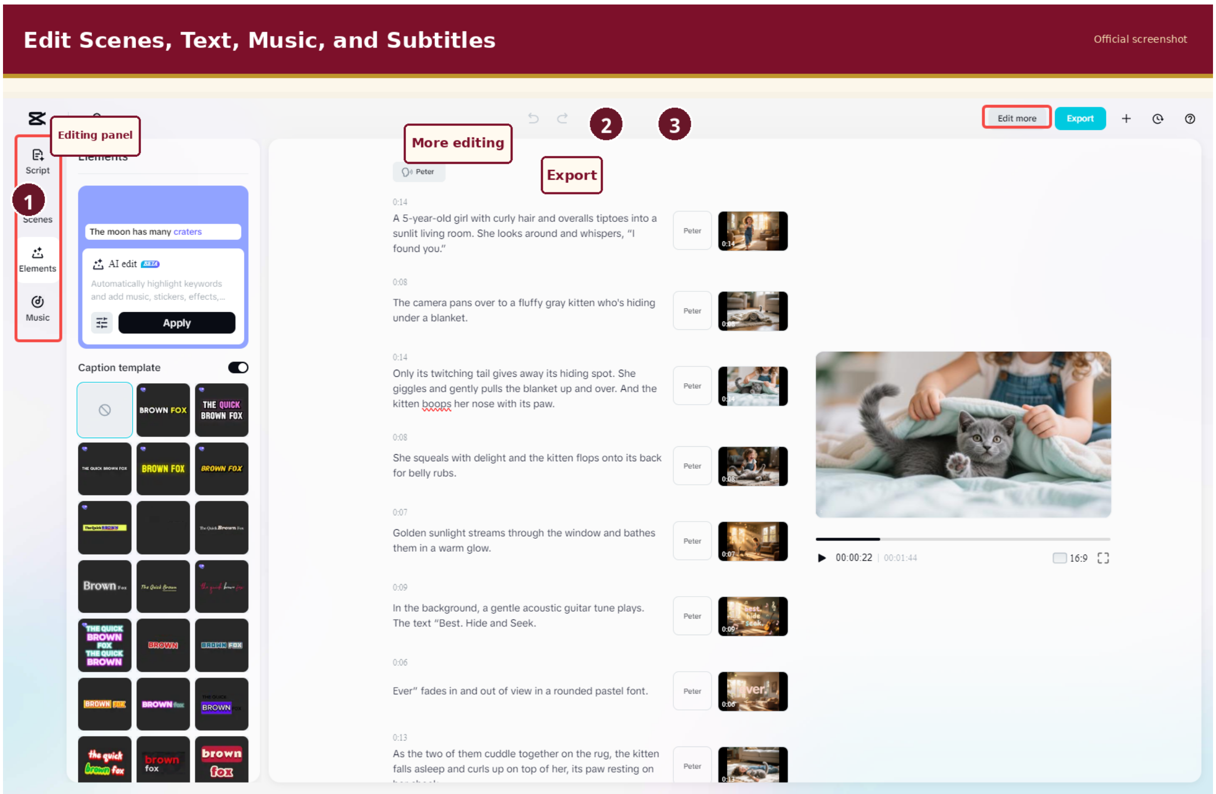


Figure 9.3 — Editing Scenes, Text, Elements, and Music

Do Not Trust Automatically Selected Visuals Blindly

If the script contains the word “Panji,” the tool may choose a generic register, a wedding scene, or an unrelated office. For local subjects, replace such visuals with your own images, archival material, maps, and specially generated clips.

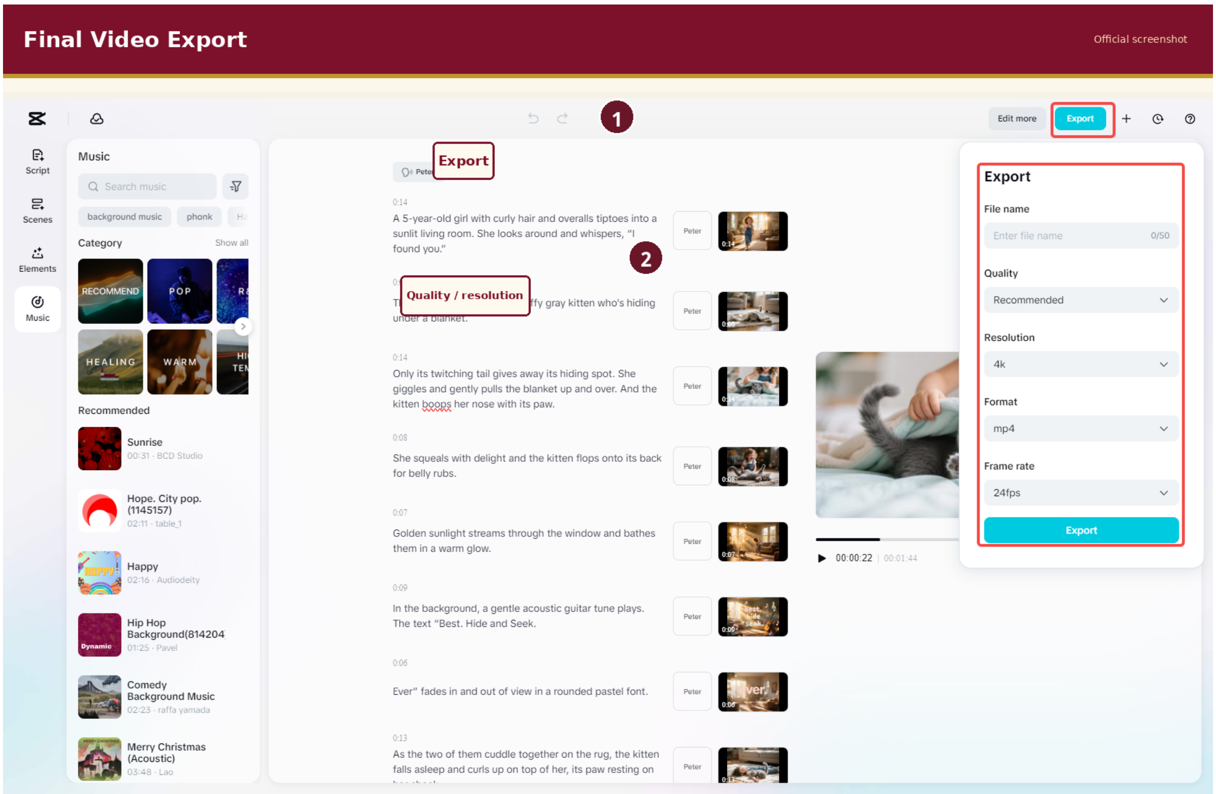


Figure 9.4 — Final Export

Common Export Settings

Use	Aspect Ratio	Resolution	Recommendation
Standard YouTube	16:9	1920×1080	MP4, original frame rate
Shorts/Reels	9:16	1080×1920	Keep main text within the safe central area
Archival copy	Original	Highest useful	Low compression; store separately
Preview	Target ratio	720p	For quick review

Voice, Music, Sound, and Dubbing

After this chapter, you will be able to —

- prepare natural Maithili narration
- choose between TTS and your own recording
- balance music and sound
- preserve consent in dubbing

Key terms: TTS · voice cloning · dubbing · sound effects · mixing

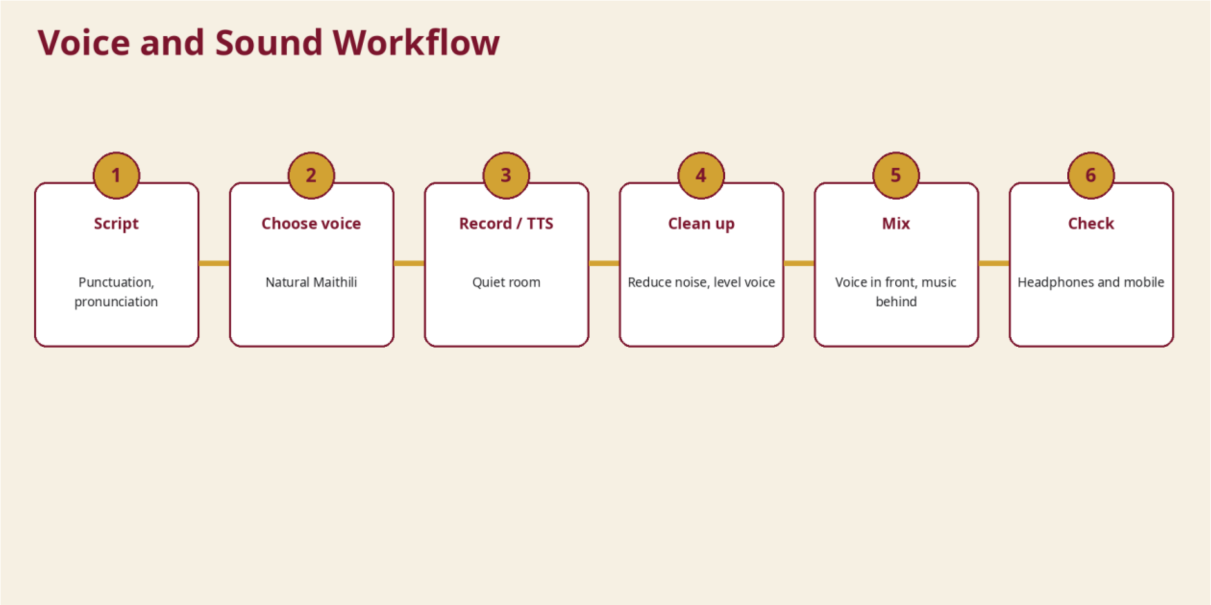


Figure 10.1 — Voice and Sound Workflow

Original teaching illustration

Three Options

Record your own voice: this gives the greatest control over Maithili pronunciation and emotion; even an ordinary microphone is sufficient if the room is quiet.

Text-to-speech: fast and useful, but limitations may remain depending on the language list, pronunciation, rhythm, and plan.

Voice conversion/cloning: use only your own voice or a voice for which clear consent has been obtained; do not impersonate others.

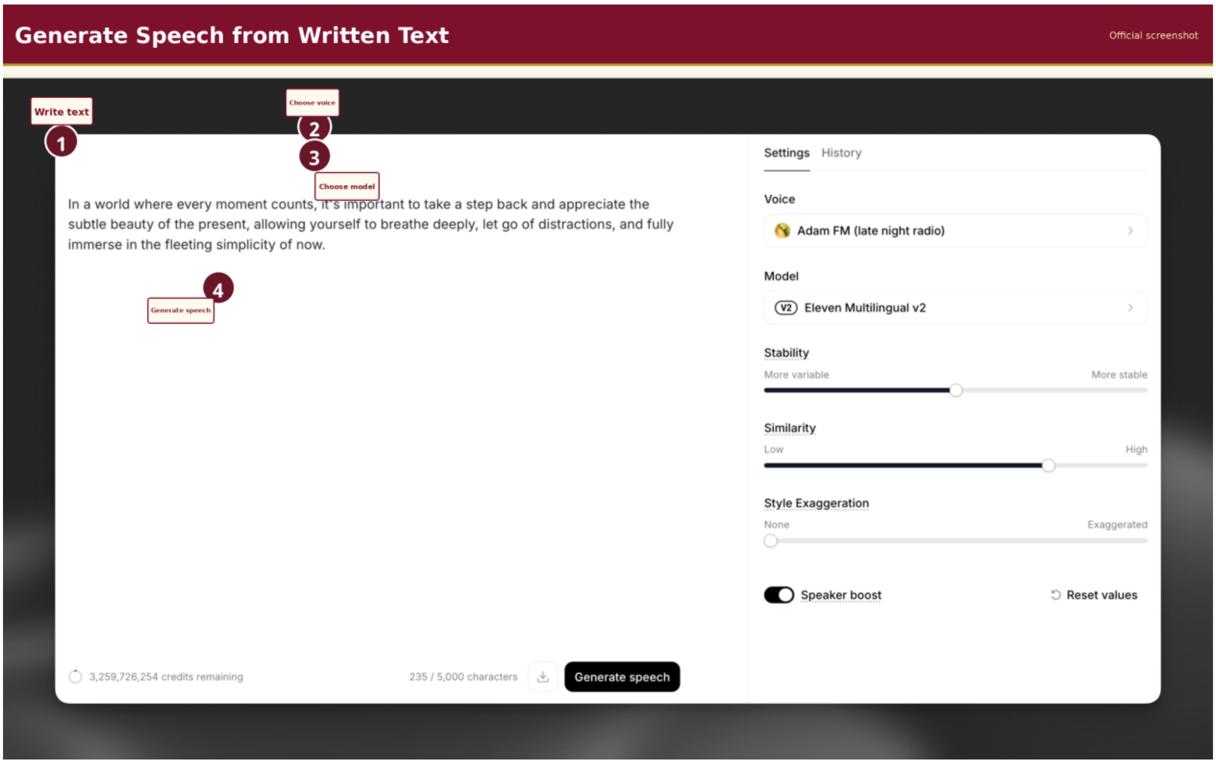


Figure 10.2 — Text, Voice, Model, and Generation

In a text-to-speech system, you select the written text, an appropriate voice, the language or accent, and generation settings. For Maithili, test two or three sentences before generating the complete script; listen yourself and correct pronunciation, vowel length, names, idioms, and pauses. Whether the voice is a human recording or synthetic, the principles of clarity, rhythm, audio balance, and final mixing remain equally important (Rose, 2015; Chattopadhyay, 2024).

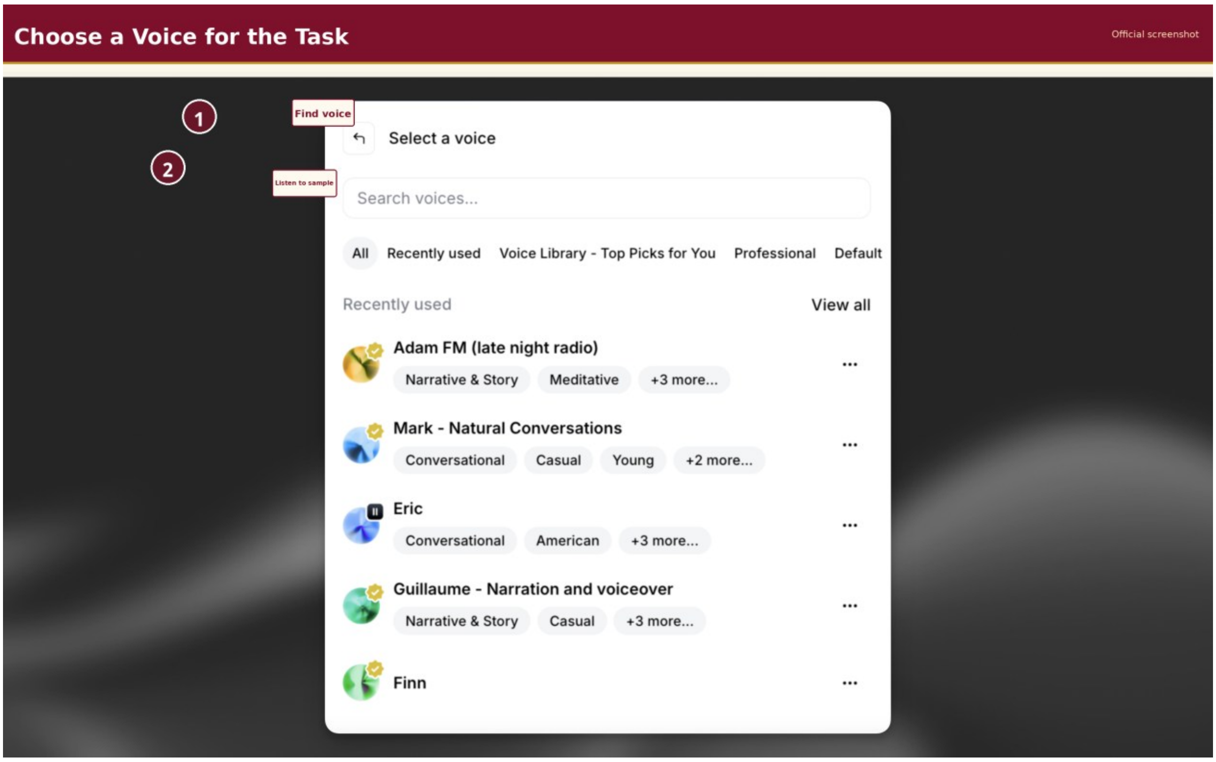


Figure 10.3 — Search for Voices by Task and Listen to Samples

Improving Maithili Pronunciation

- Write a phonetic form of difficult names separately and test it.
- Avoid very long paragraphs; divide the text into units of two to four sentences.
- Use commas, full stops, longer pauses, and line breaks to control rhythm.
- Write foreign abbreviations or numbers in the form in which they should be spoken.
- Listen to the final voice and manually correct mispronounced words; do not approve it merely by looking at the waveform.

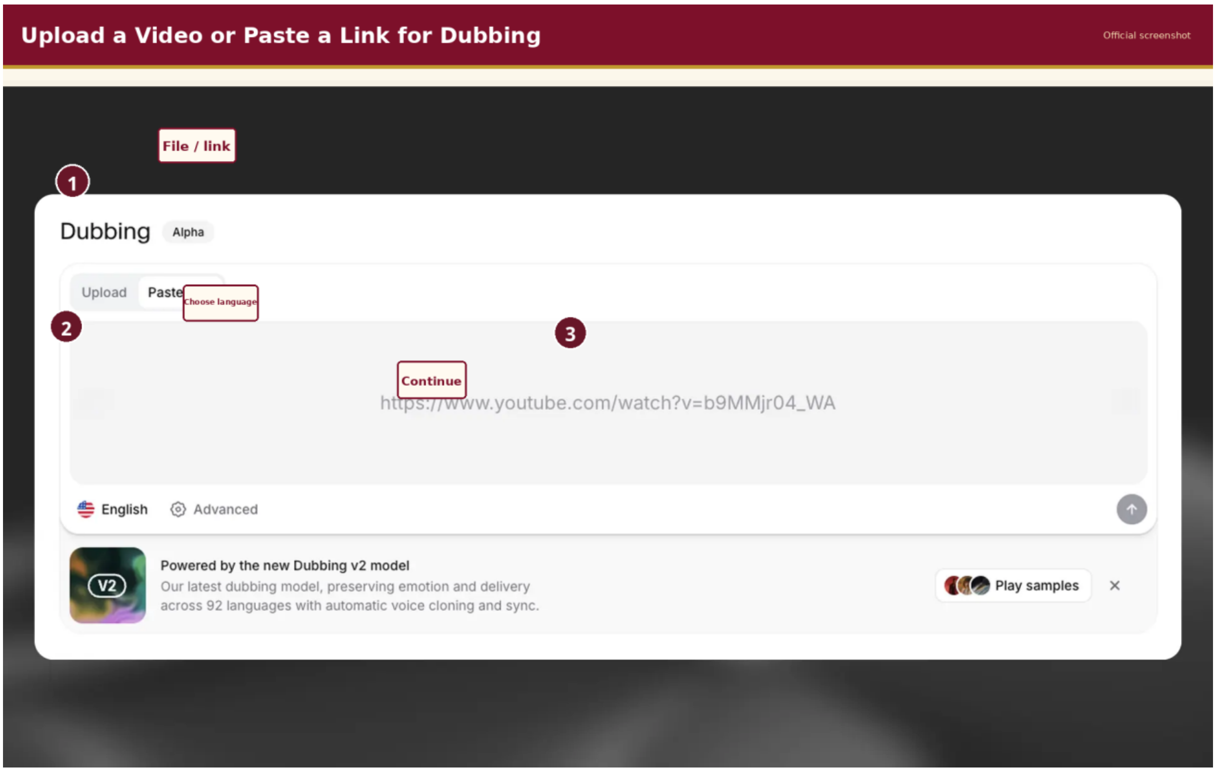


Figure 10.4 — Dubbing Project from a File or Link

An automated dubbing workflow includes the original audio or video, target language, translated text, voice similarity, timing alignment, and final mix. Automatic translation may distort Maithili meaning, forms of address, gender, idioms, and cultural context; therefore, human-edit the dubbing script before use.

Music Levels

Narration is primary. Keep the music low enough for every consonant to remain clear. It is not necessary to change the music every time the mood shifts. Use copyright-free or properly licensed music and keep a copy of the licence in the 09_sources folder.

> Consent: obtain written permission before imitating anyone's voice. A voice resembling a deceased person, public figure, or family member can also create confusion, disrespect, or legal disputes.

Editing: From Timeline to Final Cut

> After this chapter, you will be able to —
>
> • remove unused clips
>
> • improve rhythm, cuts, transitions, and colour
>
> • synchronize voice, music, and visuals
>
> Key terms: timeline · cut · J-cut · L-cut · transition · colour matching

Editing Is the Art of Selection

You may generate ten AI clips, yet only the three best may belong in the final video. Keeping a weak clip merely because time was spent generating it is a “sunk cost” error. The viewer’s experience matters, not the effort spent on generation.

First Rough Cut

107. Place the voice-over or a temporary narration on the timeline in script order.
108. Place the relevant visual above each sentence.
109. Trim unstable frames from the beginning and end of each shot.
110. Remove repetitive or weak visuals.
111. Watch the complete video once without stopping.

Cuts and Transitions

A simple straight cut is best in most situations. Use a fade at the beginning, end, or for a change of time; use a dissolve for memory or a slow transition. A different flashy transition on every scene can make a book introduction or educational video look cheap.

J-Cuts and L-Cuts

In a J-cut, the audio of the next scene begins before the picture changes. In an L-cut, the audio of the previous scene continues briefly over the next picture. These techniques make dialogue, interviews, and documentaries flow more naturally.

Colour Matching

- Keep the brightness of all clips approximately consistent.
- If one scene is excessively blue and another excessively yellow, balance the colour temperature.
- Skin tones should not look unnatural.
- Preserve detail in dark areas; do not let bright areas burn out to featureless white.
- Do not add excessive sharpening to AI-generated visuals.

Exercise 11.1 — A Thirty-Second Edit

112. Place five clips and one voice file on the timeline.
113. Remove at least one clip completely.
114. Use straight cuts in three places and a fade in one place.
115. Lower the music enough for the voice to remain clear.
116. Watch and listen on both a phone and headphones.

Exercise output: a clean, export-ready 30-second sequence.

Subtitles, Accessibility, and Maithili Script

After this chapter, you will be able to —

• correct automatically generated subtitles

• choose Devanagari fonts and line breaks

• make video useful for viewers watching without sound

Key terms: subtitles · SRT · Devanagari · safe area · alternative text

Subtitles Are Not Mere Decoration

Many viewers watch videos on phones with the sound turned off. Subtitles are essential for deaf or hard-of-hearing viewers, people in noisy places, language learners, and anyone trying to understand unfamiliar names.

Rules for Maithili Subtitles

- Use no more than two lines at a time.
- Keep each line to about 30–42 characters; test it on a phone.
- Break lines at meaningful points in the sentence, not in the middle of words.
- Use light text with a dark background or shadow; do not rely on colour alone.
- Use a clear Unicode font such as Noto Sans Devanagari.
- Check automatically generated names, anusvara, chandrabindu, conjunct letters, and Maithili verb forms yourself.

Where to Place Text

In a 9:16 video, platform buttons generally appear on the right and near the bottom. Do not place the main subtitle at the very bottom; keep it within a "safe central area." Do not cover a face or the main object with text.

SRT and Burned-In Subtitles

SRT: a separate text file; viewers can turn it on or off, and it assists search and translation.

Burned-in subtitles: permanently embedded in the image; visible on every platform, but harder to correct later.

Quality check: do not judge subtitle accuracy only by reading the text. Play the video and check whether subtitle timing matches the spoken audio; timing is equally important.

YouTube, Shorts, and Social Publishing

After this chapter, you will be able to —

• write an effective title, thumbnail, and description

• disclose AI-generated content

• adapt one video for several platforms

Key terms: thumbnail · description · disclosure · metadata · C2PA

Final Package Before Publication

- Final MP4

- Thumbnail
- Three title options
- A two-paragraph description
- List of sources and licences
- Subtitle file
- Disclosure of AI use
- Relevant link or next lesson

Title

A title should make a clear promise, not deceive. “A Three-Minute Maithili Book-Introduction Video with AI” is clear. A misleading title such as “You Will Be Shocked by This Secret” may produce short-term clicks but not trust.

Thumbnail

- Use one main face or object.
- Use 3–6 large words; add them in a design tool, not in the video generator.
- Test the thumbnail at mobile size.
- Do not show an event or person that does not appear in the video.
- Keep brand colours limited and consistent.

AI Disclosure

Realistic synthetic or altered visuals and voices can mislead viewers. Therefore, clearly disclose artificial reconstruction or alteration of real people, events, or places; state in the title, description, or opening notice which parts are synthetic. Ordinary colour correction, noise reduction, or obviously fictional animation may be different in nature; the underlying principles are transparency, consent, and accountability (Coeckelbergh, 2020; Dubber et al., 2020).

Sample clear disclosure: “Some visuals in this video were created with artificial-intelligence-based tools. The script, fact-checking, Maithili narration, and final editing were completed by a human creator.”

One Piece of Content, Three Formats

Format	Duration	Aspect Ratio	Purpose
Main video	3–8 minutes	16:9	Complete story/lesson
Short	30–60 seconds	9:16	One question or main fact
Teaser	10–20 seconds	9:16 / 1:1	Lead viewers to the main video

Copyright, Consent, Deepfakes, and Safety

After this chapter, you will be able to —

• choose rights-cleared material
• understand consent for faces and voices
• avoid deceptive realistic synthetic scenes
• understand the broad direction of Indian rules
Key terms: copyright · licence · consent · deepfake · synthetic information

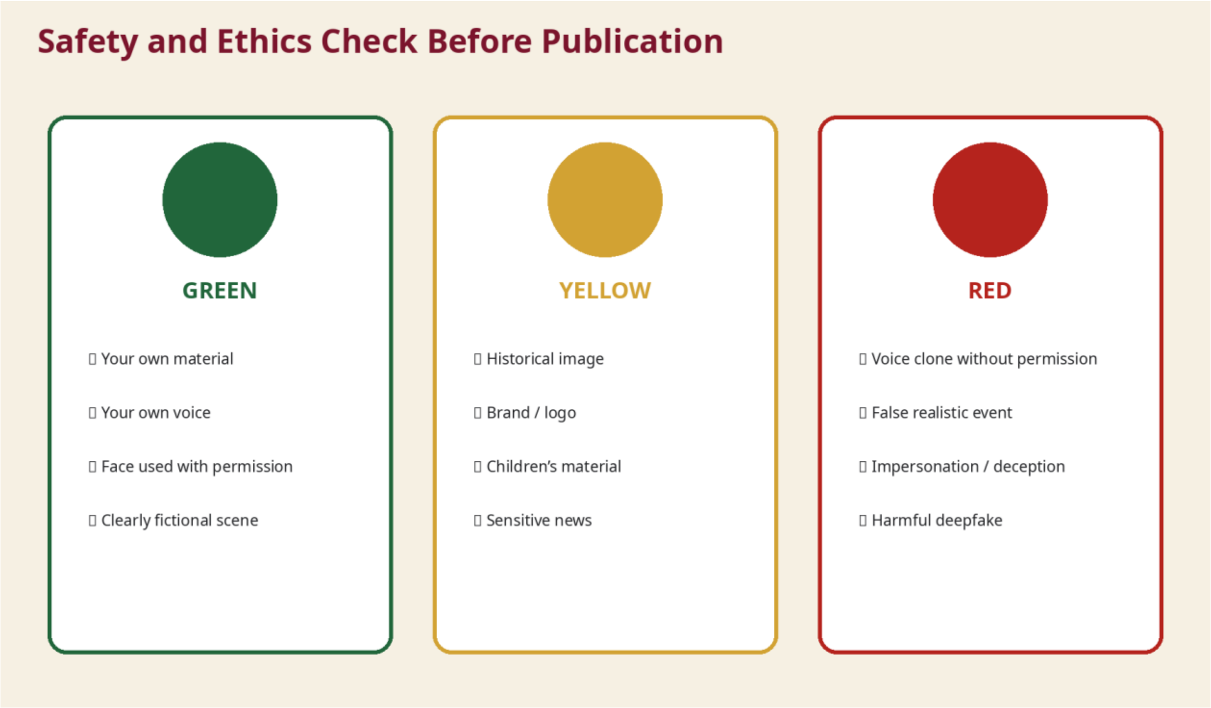


Figure 14.1 — Ethical and Safety Signals

Original teaching illustration

Four Questions

117. Who owns this image, video, music, or voice?
118. Do I have permission to use, modify, and publish it?
119. Could viewers mistake it for something real?
120. Could this material cause unfair harm to any person, community, or institution?

Keep Proof of Rights

For every external asset, save the link, author, licence, download date, and permission in the 09_sources folder. “Found on the internet” does not mean “free to use.” The terms of service of AI tools also differ; read them before commercial use.

Faces and Voices

Using your own photograph or voice is straightforward, but for another person’s face, private photograph, or voice, treat explicit, informed, written consent as necessary. Permission merely to “use a photograph” does not automatically include AI transformation or voice cloning. For children’s material, parental permission and special caution are required.

Synthetic Content in India

When publishing synthetic or altered realistic content, consider risks involving consent, privacy, defamation, copyright, personality rights, and misleading presentation. Rules may differ by country, platform, and context; adopt clear labelling, source verification, and responsible publishing practices, and seek qualified legal advice for disputed or commercial uses (Dubber et al., 2020; Liao, 2020).

What Not to Do

- Do not make someone appear to say something they never said.
- Do not create false realistic scenes concerning disasters, elections, war, illness, or finance.
- Do not imitate someone else's voice for phone calls, advertisements, endorsements, or jokes.
- Do not change the date or location of archival images and present them as factual.
- Do not evade responsibility for a visible AI error by saying, "the tool made it."

Legal boundary: this chapter provides general digital-literacy guidance, not individualized legal advice. Laws and platform rules continue to change.

Three Complete Practical Projects

After this chapter, you will be able to —

• apply the entire process from beginning to end

• create versions for different durations and platforms

• produce original Maithili content

Key terms: short Reel · educational video · book introduction · production sheet

Project 1 — 30-Second Visual Poem: "Dawn at the Pond"

Objective: a 9:16 mobile Reel, five scenes, your own or rights-cleared images, and poetic narration in Maithili.

- Scene 1: the first light of dawn settles on the pond water.
- Scene 2: among the makhana leaves, the breeze writes its delicate letters.
- Scene 3: one stroke of a distant boat makes circles of time upon the water.
- Scene 4: the village has not yet awakened, but the birds are already announcing the day's first news.
- Scene 5: this pond is not merely water — it is a mirror of memory, livelihood, and life.

Visual prompt: Vertical 9:16, a calm Mithila village pond at dawn, makhana leaves, soft mist, subtle ripples, distant wooden boat, slow cinematic push-in, natural documentary colors, peaceful mood, no text, no logo.

Production Sequence

121. Choose a 9:16 aspect ratio.
122. Create one wide shot, two close-ups, one boat shot, and one closing shot.

123. Record your own voice in a quiet room.
124. Add soft natural sound.
125. Keep subtitles to no more than two lines.
126. Add an AI-visual disclosure in the description.

Exercise output: a 30-second MP4, thumbnail, and SRT.

Project 2 — Three-Minute Educational Video: "How to Read a Maithili Book"

Objective: a 16:9 explanatory video using screen recording, book pages, simple animated cues, and clear subtitles.

Time	Topic	Visual
0:00–0:15	Hook: check three things before opening a book.	Book, notes, simple icons, presenter or screen
0:15–0:50	Author, genre, and publication context.	Book, notes, simple icons, presenter or screen
0:50–1:30	First reading: understand the story or argument.	Book, notes, simple icons, presenter or screen
1:30–2:10	Second reading: note language, symbols, and structure.	Book, notes, simple icons, presenter or screen
2:10–2:40	Note questions and quotations.	Book, notes, simple icons, presenter or screen
2:40–3:00	Summary and next book.	Book, notes, simple icons, presenter or screen

Visual principle: an educational video does not need a new AI-generated image for every sentence. Real book pages, screen recording, simple line graphics, title cards, and limited B-roll are often more trustworthy.

Project 3 — Five-Minute Book Introduction

Objective: do not reveal the whole story, but clearly explain the subject, style, historical importance, intended reader, and benefit to the reader.

- 0–20 seconds: the book's central question.
- 20–60 seconds: the author and context.
- 1–2 minutes: subject and structure.
- 2–3 minutes: three distinctive features.
- 3–4 minutes: one short quotation or visual example — with proper rights.
- 4:00–4:40: who should read it and why.
- 4:40–5:00: sources, book link, and conclusion.

Project Check

- Are the story or book facts accurate?
- Are the rights to use the cover and quotations clear?
- Are the generative visuals clearly presented so they cannot be mistaken for authentic illustrations from the book?
- Is the voice natural Maithili?
- Is there no unnecessary repetition within the five minutes?
- Does the description provide the book, author, publisher, and sources?

Common Problems and Solutions

After this chapter, you will be able to —

• identify defects in visuals, voice, text, and export

• apply targeted corrections

• decide when to regenerate and when to edit

Key terms: defect diagnosis · regeneration · mask · noise · flicker

Defect	First Solution
Face keeps changing	Use the same reference image, stability instructions, and shorter clips; cut the failed section.
Hands are distorted	Keep hands outside the frame, simplify the action, avoid extreme close-ups; choose another clip.
Clothing changes	Repeat the same character-sheet description; use image-to-video.
Camera moves too fast	Use “locked camera” or “very slow push-in”; increase duration if appropriate.
Scene flickers	Reduce motion, use a stable background, shorten the clip; use stabilization sparingly in the editor.
Generated text is wrong	Remove generated lettering from the scene; add Unicode text in the editor.
Maithili voice is incorrect	Shorten sentences, use phonetic guidance, record your own voice, or make a human correction.
Music masks the voice	Lower music by roughly 10–18 dB; use automatic ducking during speech.
Subtitles are cropped	Use the safe area, avoid tiny fonts, and preview in 9:16.
File is too large	Use H.264 MP4, a sensible bitrate, and avoid unnecessary 4K.

Regenerate or Edit?

Regenerate: when the main face, body, object, or action is fundamentally wrong.

Edit: when the beginning or end is poor, the duration is too long, the colour differs slightly, or there is a minor background defect.

Mask/crop: when the defect appears for only a moment or in a small part of the frame.

Discard: when the clip's meaning or ethics are wrong and correction would take too long.

Golden rule: do not regenerate until you can name the defect. "It does not look good" is not an actionable diagnosis.

Twenty-One-Day Self-Study Programme

After this chapter, you will be able to —

• build skill through regular practice

• prepare three complete samples

• write down your own production system

Key terms: daily practice · portfolio · review · publication

Day	Task	Done
Day 1	Understand AI video types; identify five examples.	☐
Day 2	Create folders; make four aspect-ratio samples.	☐
Day 3	Write a one-minute idea and define the audience.	☐
Day 4	Write a 60–90 second script.	☐
Day 5	Create a six-scene storyboard.	☐
Day 6	Write five seven-element prompts.	☐
Day 7	Make three text-to-video tests.	☐
Day 8	Create three motion variants from one image.	☐
Day 9	Create a character sheet and continuity test.	☐
Day 10	Record your own voice.	☐
Day 11	Make a short TTS comparison.	☐
Day 12	Clean audio and mix music.	☐

Day 13	Create a rough timeline edit.	☐
Day 14	Refine cuts, transitions, and colour matching.	☐
Day 15	Create Maithili subtitles.	☐
Day 16	Complete the 30-second project.	☐
Day 17	Write the script and storyboard for a 3-minute project.	☐
Day 18	Create and edit the visuals.	☐
Day 19	Prepare the publication package, thumbnail, and description.	☐
Day 20	Check copyright, consent, and AI disclosure.	☐
Day 21	Publish the final video, complete self-evaluation, and set the next goal.	☐

Final Self-Evaluation

- ☐ Topic and audience are clear
- ☐ Script fits the allotted time
- ☐ Visual continuity is consistent
- ☐ Prompts have been saved
- ☐ Voice sounds natural
- ☐ Music is balanced
- ☐ Subtitles are accurate
- ☐ Rights/consent are clear
- ☐ AI disclosure is appropriate
- ☐ Final file is backed up

Next goal: after twenty-one days, choose one genre — book introduction, educational lesson, folktale, archival documentary, or short social video — and develop a consistent visual grammar across ten videos.

Appendix 1 — Ready-to-Use Prompt Collection

The prompts below are starting templates. Change the person, place, time, aspect ratio, and style according to your project. Before publishing a realistic synthetic scene of a real person or place, check consent and disclosure requirements.

1. Dawn in a Mithila Village

Prompt: A wide documentary shot of a Mithila village at dawn, mud courtyard, tulsi altar, soft

mist, a woman drawing aripan, very slow camera push-in, natural colors, respectful realistic style, no text, no logo.

2. Pond and Makhana

Prompt: Aerial wide shot of a calm pond filled with makhana leaves, gentle morning breeze, small wooden boat moving slowly, soft golden sunlight, realistic documentary style, stable camera.

3. Subtle Motion in a Madhubani Painting

Prompt: Animate the original Madhubani painting while preserving flat colors, linework and composition; fish tails move gently, lotus leaves shimmer, no 3D transformation, slow camera push-in.

4. Opening a Book

Prompt: Close-up of an old Maithili book opening on a wooden desk, pages turning slowly in a soft breeze, warm window light, shallow depth of field, cinematic but realistic, no readable generated text.

5. Writing Scene

Prompt: Top-down shot of a hand writing notes beside an open book, slow natural hand movement, warm study lamp, stable camera, documentary educational mood.

6. Classroom Lesson

Prompt: Bright classroom in rural Mithila, teacher pointing to a simple diagram, attentive students, medium shot, natural daylight, subtle camera movement, realistic educational documentary.

7. Archival Atmosphere

Prompt: Slow pan across carefully preserved old documents and family registers on a clean table, museum lighting, respectful archival documentary style, no invented readable text.

8. Product Introduction

Prompt: Clean studio shot of a handmade Mithila craft object rotating slowly on a neutral background, soft diffused light, accurate shape and color, commercial product video, no text.

9. Rain

Prompt: A locked-camera shot of monsoon rain falling in a Mithila courtyard, water drops

from tiled roof, banana leaves moving gently, realistic ambience, soft gray light.

10. Train Journey

Prompt: View through a train window moving across green fields in North Bihar, light rain on glass, calm travel documentary mood, stable realistic motion.

11. Folktale Forest

Prompt: Stylized 2D illustrated forest inspired by Mithila line art, moonlight, leaves moving softly, a deer pauses and listens, slow side tracking, clearly fictional animation.

12. Sunrise Time-Lapse

Prompt: Wide static shot of sunrise over fields, clouds and light changing gradually, smooth time-lapse feeling, natural colors, no sudden camera movement.

13. Mobile Short Hook

Prompt: Vertical 9:16 macro close-up of makhana popping open during roasting, quick but clear action, dramatic natural sound, high detail, no text.

14. Interview B-Roll

Prompt: Medium close-up of an empty chair, microphone and books prepared for an interview, soft window light, very slow push-in, professional documentary mood.

15. Publication Closing Shot

Prompt: A book resting beside a cup of tea and a lit lamp, evening ambience, slow pull-back revealing a quiet study, warm natural colors, space for title added later in editor.

16. Motion in a Still Image

Prompt: Keep the person’s face, clothing and background unchanged. Only the eyes blink naturally, the fabric moves slightly in the breeze, and the camera makes a very slow push-in.

17. Reflection on Water

Prompt: Close-up of water reflections with lotus leaves, gentle ripples spreading from a single drop, locked camera, soft morning light, meditative realistic mood.

18. Artisan’s Hands

Prompt: Close-up documentary shot of artisan hands painting fine lines on a handmade object,

accurate finger movement, stable camera, soft side light, no face shown.

19. Map Alternative

Prompt: Minimal animated diagram showing a route with clean lines and simple icons on a plain background; no generated labels, leave empty spaces for text added in editing.

20. Title Background

Prompt: Subtle cream handmade paper texture with a thin burgundy Mithila-inspired border, very gentle light movement, empty center for title, no generated text.

Appendix 2 — Worksheets and Checklists

A. Project Sheet

Topic	Write
Project name	
Audience	
Main objective	
Duration	
Platform	
Aspect ratio	
Main message	
Required sources	
Rights/consent	
Final publication date	

B. Shot List

No.	Time	Voice	Visual	Prompt/Source	Status

C. Pre-Publication Check

- ☐ Title is truthful and clear
- ☐ Thumbnail matches the video
- ☐ Voice is clear
- ☐ Music is balanced
- ☐ Maithili subtitles are accurate
- ☐ Sources for factual claims are recorded
- ☐ Image/music licences are documented
- ☐ Consent for faces/voices is documented
- ☐ AI use is disclosed
- ☐ Final file and project are backed up

Appendix 3 — Video Glossary

Term	Meaning
Aspect ratio	Relationship between image width and height; for example, 16:9.
B-roll	Supporting visuals shown over the main narration or interview.
Clip	A short segment of video.
Cut	A direct change from one shot to another.
Dolly	Physical movement of the camera forward or backward.
FPS	Frames per second; influences how smooth motion appears.
Flicker	Unwanted fluctuation in light, texture, or objects.
Frame	A single still image within a video.
Hook	The opening element that captures the viewer’s attention.
Image-to-video	A generative process that adds motion to a still image.
J-cut	The audio of the next scene begins before the picture changes.
L-cut	The audio of the previous scene continues

	over the next picture.
Lip-sync	Matching mouth movement with spoken audio.
Mask	A selected area of an image to which a separate edit is applied.
Pan	Horizontal rotation of the camera from left to right or right to left.
Prompt	A written or visual instruction given to an AI tool.
Resolution	The number of pixels in an image; for example, 1920×1080.
Reference image	An image supplied to preserve appearance, character, or composition.
Seed	A numerical value that can influence repeatability in some tools.
Storyboard	A scene-by-scene visual plan for a video.
SRT	A subtitle text file containing time codes.
Text-to-video	A process for generating new video clips from written descriptions.
Timeline	The time-ordered area in an editor where video, voice, and text are arranged.
Tilt	Vertical camera rotation up or down.
Transition	An effect used to move between two scenes.
TTS	Converting written text into synthetic speech.
Upscale	Increase the resolution of video or an image.
Voice clone	A synthetic voice designed to resemble someone’s voice; explicit consent is required.

Part 2
Expanded Explanation and Application

This book is for everyone who wants to communicate through video but has neither a camera, nor a studio, nor training in editing. Artificial intelligence now makes all three tasks possible with an ordinary mobile phone or computer. Work that once required equipment costing hundreds of thousands and a team of many people can now begin with a written instruction — a "prompt."

But let one thing be clear from the very beginning. AI is not a magic wand. It is a tool — like an axe, like a pen. An axe cuts wood, but the carpenter decides which tree to cut and what house to build. In the same way, AI can create visuals and voices, but what to say, why to say it, and whom to say it to — the answers to all three questions remain with you. This book will teach you to operate the tools; the craft will remain your own.

This book is written in a "Teach Yourself" style — that is, no teacher or training centre is required. Each chapter is one step. Begin with the first chapter, proceed in order, and be sure to complete the "Exercise" at the end of each chapter. Reading and doing must go together; only then will learning take place.

One more point: the illustrations in this book are original teaching illustrations, not copies of the actual screens of any company. There are two reasons. First, AI-tool interfaces change every few months, so real screenshots may already be outdated by the time they are printed. Second, once you understand the principle, no new screen will feel unfamiliar; if you merely memorize screens, every change will stop you. The illustrations in this book show the underlying patterns shared by many tools — where prompts are entered, what kinds of settings appear, and what a timeline looks like. Once you understand this structure, you will be able to open a new tool and learn to use it yourself — and that is the true meaning of "Teach Yourself."

The shortage of technical books in Maithili has long been felt. My request is that whatever you learn from this book, use it to make videos in Maithili and show them to the world. The stories, songs, paintings, history, festivals, and traditions of Mithila are all waiting to be seen in video.

What Is AI Video?

Artificial intelligence (AI) is the ability of computers to perform tasks associated with human intelligence — understanding language, recognizing images, and now creating images and video. When we write to an AI system, "Show two children playing in a pond," and it creates a moving scene of that kind, this is called "generative AI." To generate means to produce. The AI is not searching somewhere for an existing video and bringing it back; it is creating a new video that did not previously exist in that form.

How is this possible? In simple terms, these AI models have been trained on enormous numbers of images and videos. From this exposure they learn what an egret looks like, how water moves, what colour dawn light can be, and how a walking person's legs move. When you write "an egret beside a pond at dawn," the model constructs a scene from the patterns it has learned. It is somewhat like a painter who has seen thousands of ponds and can later paint one from memory without having a pond directly in front of them.

There are mainly four ways to make video with AI:

First — text-to-video. You write a description, and the AI creates the whole scene. This is the most striking method, but it offers somewhat less control — the result may not match your imagination exactly.

Second — image-to-video. You provide a still image — a photograph or an AI-generated picture — and the AI makes it move. This gives greater control because you have already selected the starting scene.

Third — avatar video. A digital presenter, or avatar, appears on screen and speaks your written text, like a newsreader. This is especially useful for lessons, announcements, or lectures.

Fourth — AI-assisted editing. Here, the video itself may be your own recording, while AI helps with trimming, joining, creating subtitles, or removing backgrounds.

This book covers all four paths, because in practice a good video is often made from a mixture of them.

Now understand what AI cannot do reliably. AI still does not create a long, coherent video in one pass; it usually generates short clips of roughly five to twenty seconds, which are then joined into a longer sequence. AI-generated scenes may also contain errors — the number of fingers may change, written letters may become distorted, or a character's face may change from one clip to the next. Most importantly, AI has no lived knowledge of Mithila, Maithili, or what is in your heart; it can only respond to what you know how to ask for. That is why a large part of this book is devoted to "how to ask" — in other words, planning and prompting.

Exercise: on paper, write three video topics you would like to make. For each, write one line answering two questions: who will watch this video, and what will they gain from watching it?

Getting Started: What Do You Need?

You do not need expensive equipment to make AI video. What you need is quite modest:

First, a device. A smartphone is enough; a computer or laptop is more convenient because it makes it easier to write prompts on a large screen, manage files, and edit. A particularly powerful computer is usually unnecessary because generation happens on the company's servers, not on your device. Your device sends instructions and downloads the completed video.

Second, internet access. The better the connection, the less you will wait. Video files are large, so downloading will consume data — allow for several gigabytes per month.

Third, an email account. Most AI tools require an account, and many allow direct sign-in with Google/Gmail. A practical suggestion: keep a separate email account for your creative work so that notifications from AI services do not fill your main inbox.

Fourth, understand "credits." Many AI video tools operate on a credit system. Credits are like coupons: each generation consumes some of them. Free accounts may provide a small allowance — daily, monthly, or only once. Paid plans generally provide more credits, higher quality such as 1080p or 4K, watermark-free output, and longer clips.

Adopt a simple practical policy: test cheaply first, then spend only when the tool has proved useful. While learning a new tool, use free credits for small experiments. Learn how to prompt and understand the tool's behaviour. When you are ready for serious work — for example, running a YouTube channel — subscribe to one appropriate tool. Buying plans for every tool is unnecessary; one or two are enough.

Fifth, establish a file-management system. This sounds minor, but experienced creators know how quickly work disappears without organization. Create a separate folder for each video project on your computer or phone. Inside it, keep four subfolders: “Script” (script and prompts), “Clips” (raw AI-generated video), “Voice-Music” (voice-over and music), and “Final” (finished video). Give every file a meaningful name. A file called “clip-1-pond-dawn” will still make sense six months later; “video(3)-final-new-2” will not.

And finally — patience. AI tools are sometimes busy, sometimes produce strange results, and sometimes fail to understand a good prompt. This is normal. Those who are willing to make three or four attempts learn the process; those who abandon it after one odd result miss the opportunity.

Exercise: create the folder structure described above on your device. If you wish, create a new email account. Then open the site of one AI video tool and simply examine what its free plan offers. Do not make anything yet; just observe.

Planning a Video: From Idea to Script

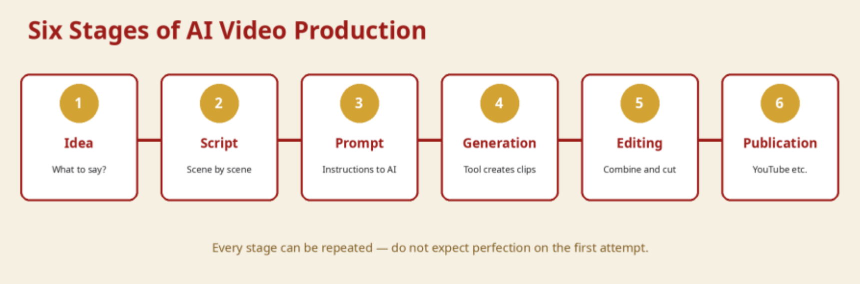


Figure 1 — Six Steps of AI Video Production

The biggest mistake beginners make is to open a tool and start pressing buttons immediately. An unplanned video looks like a house built without a blueprint. That is why this chapter comes before the tools.

Before making any video, write answers to three questions:

One — what is the purpose? To teach (for example, “how Madhubani painting is made”), to tell (a story), to show (a village scene), or to announce something? One video, one purpose — follow this rule. A video that tries to contain everything often communicates nothing clearly.

Two — who is the audience? Children, students, Maithili speakers living away from Mithila, curious outsiders, or a general audience? The audience determines how simple the language should be, how long the video should be, and how the visuals should feel. For children: brighter colours and quicker pacing. For adults: more pause and seriousness.

Three — how long? At the beginning, make videos of 30 seconds to 2 minutes. Short videos are easier to make, easier to improve, and well suited to contemporary viewing habits. Remember: if AI produces five- to ten-second clips, a one-minute video means roughly six to twelve clips.

Now, the script. A script need not be a complicated literary object. At its simplest, it is a two-column table. In the left column, write “What will be seen” (visual); in the right, “What will be heard” (voice or subtitle). Five or six lines may be enough for a thirty-second video.

AI can assist here in another role — as a scriptwriting aide. Give a chatbot such as Claude, ChatGPT, or Gemini your topic and ask it to draft a script. But you must ask intelligently. If you merely say, “Write a video script about Madhubani,” you are likely to receive generic, lifeless text. Ask more specifically:

"I am making a 60-second YouTube video about Madhubani painting. Audience: young Indians who have heard the name Madhubani but know little about it. Tone: simple and warm, in Maithili. I want a two-column script — the left column should describe the visuals that I can give to an AI video tool, and the right column should contain the voice-over. Six scenes, ten seconds each."

Notice how this specifies the subject, duration, audience, language, format, and intended use. The clearer the request, the better the result. The same principle will apply later to video prompts.

Do not accept a chatbot-written script with your eyes closed. Read it and check it against your own knowledge. Are the facts correct? Does the language sound like you? Revise lines you do not like or ask for focused changes, such as "make the third scene more emotional" or "shorten the final line." The script is your signature; AI is only a writing assistant.

Exercise: take one of the three topics selected in Chapter 1. Write answers to the three questions — purpose, audience, and duration. Then use a chatbot to generate a two-column script using a detailed request of the kind shown above, and make at least two revisions yourself.

Storyboard: A Map of the Visuals

Figure 2 — Storyboard for the "My Village" Video: Sketch, Duration, and Camera Cue for Each Scene

A script is a map of words; a storyboard is a map of visuals. In a storyboard, a rough sketch or written description of every scene is placed in sequence. Filmmakers have used this method for more than a century, and in the AI era its importance has increased rather than decreased. Why? Because AI needs a separate, clear instruction for every clip, and the storyboard becomes that instruction list.

Do not worry — you do not need to know how to draw. A round face, stick-like body, and arrows are enough. If you do not want to sketch at all, write three lines for each scene: what is visible, what the camera does, and how many seconds the shot lasts.

Learn a little camera language, because these terms will be useful later in prompts:

Wide shot — the whole scene from a distance; used to establish a place. Example: a full view of a village.

Close-up — viewed from nearby; used for emotion and detail. Example: steam rising from tea on a stove.

Tracking shot — the camera moves along with the character. Example: following a child on the way to school.

Drone shot — viewed from above, as a bird might see it. Example: an aerial view of a pond and fields.

Dolly in/out — the camera slowly moves closer or farther away; useful for emotional emphasis.

Static shot — the camera stays in one place; useful for calm, composed scenes.

Remember a simple formula for visual sequence: "establish, develop, conclude." The first scene tells us where we are — the establishment, often a wide shot. The middle scenes show the subject more closely — development, using close-ups or tracking shots. The final scene gathers the meaning — an emotion, a message, or a title card.

Figure 2 presents a storyboard for a thirty-second video titled "My Village." Notice how six scenes travel from dawn to evening — morning mist for the establishment, tea–school–pond for development, and lamp plus title for the conclusion. Each scene includes duration and a camera cue. Later, each of these scenes will become an individual prompt.

While making a storyboard, pay attention to continuity. If the first scene is at dawn, the next should not suddenly become night. If a character wears a red kurta in one scene, the red kurta should remain in every scene — and you may need to repeat that detail in every prompt because AI does not reliably remember the previous clip. Write the character description separately on a "character sheet" and paste it unchanged into every relevant prompt. This small technique is one of the simplest ways to bring consistency across clips.

Exercise: convert your script into a six-scene storyboard. For each scene, write the visual description, camera instruction, and duration. If there is a recurring character, create a three-line "character sheet" describing clothing, age/state, and appearance.

The Art of Prompt Writing

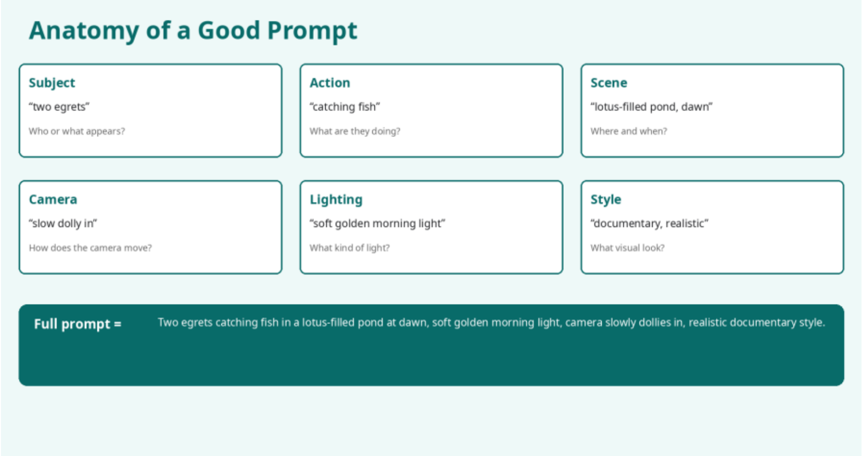


Figure 3 — Six Elements of a Good Prompt: Subject, Action, Scene, Camera, Lighting, Style

A prompt is the written instruction you give to AI. It is one of the most valuable skills of the AI era, and fortunately it does not require technical knowledge so much as clarity of language — and language is our own field.

Compare a poor prompt with a good one. Poor: "village scene." This tells the AI almost nothing — which country's village? Which season? Day or night? What is happening? The model will fill the gaps itself, often with a generic or unrelated village. Better: "A village in North India at dawn, light mist over the fields, thatched and tiled-roof houses, a banyan tree in the distance, camera

slowly panning right, soft golden light, realistic documentary style.” Now the AI has a complete visual direction.

Remember the six-element formula shown in Figure 3:

Subject — who or what is the main focus? (two egrets)

Action — what are they doing? (catching fish)

Scene — where and when? (a lotus-filled pond at dawn)

Camera — what does the camera do? (moves slowly forward)

Lighting — what kind of light? (golden dawn light)

Style — what kind of visual treatment? (documentary, realistic)

Try to include all six elements in every prompt, usually in this order — subject and action first, style toward the end. Useful style terms include realistic, cinematic, animation, watercolor illustration, vintage-film look, and documentary.

A note on language. Most AI video tools understand English prompts best, and some also understand Hindi. A practical method is therefore to do your scripting and thinking in Maithili, then prepare the final prompt in English. A chatbot can help: give it your Maithili scene description and ask, “Turn this into an English video prompt containing six elements — subject, action, scene, camera, lighting, and style.” In this way, the imagination remains Maithili; only the linguistic conversion is assisted.

Negative prompts are also common — instructions that state what you do not want. For example: “no blur, no deformed hands, no written text on screen, no watermark.” Some tools provide a separate field for negative prompts; in others, these instructions are added to the main prompt.

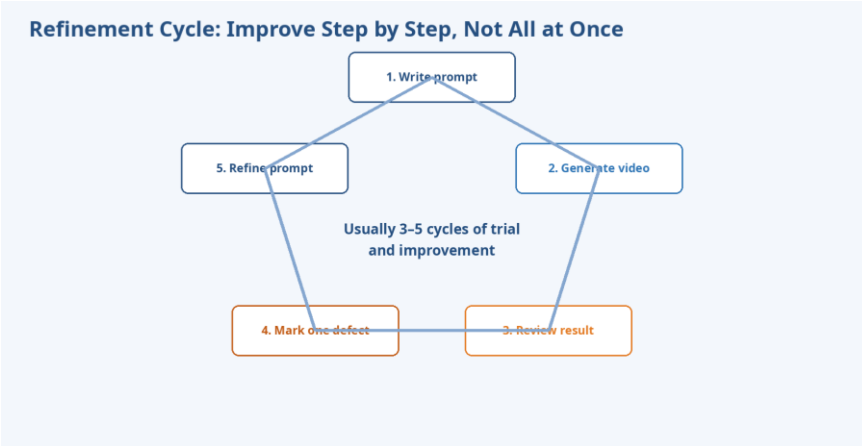


Figure 4 — The Refinement Cycle: Prompt, Result, Review, Improvement

Now the most important principle: the refinement cycle shown in Figure 4. It is rare for the first prompt to produce exactly the video you want, and that is not failure — it is the method. Watch the result, identify what is wrong, change only that part of the prompt, and generate again. Are the egrets too large? Add “small egrets.” Is the scene too bright and glossy? Change the lighting to “soft, subdued light.” Usually, three to five cycles can produce a usable clip. Change only one or two things in each cycle; if you alter everything at once, you will not know which change actually helped.

Keep a file of your successful prompts. Six months from now, your future self will thank you.

Exercise: for the first scene of your storyboard, write one complete prompt in Maithili using the six-element formula. Then prepare its English version yourself or with a chatbot. Save both in your “Script” folder.

Text-to-Video Tools: Making the First Clip

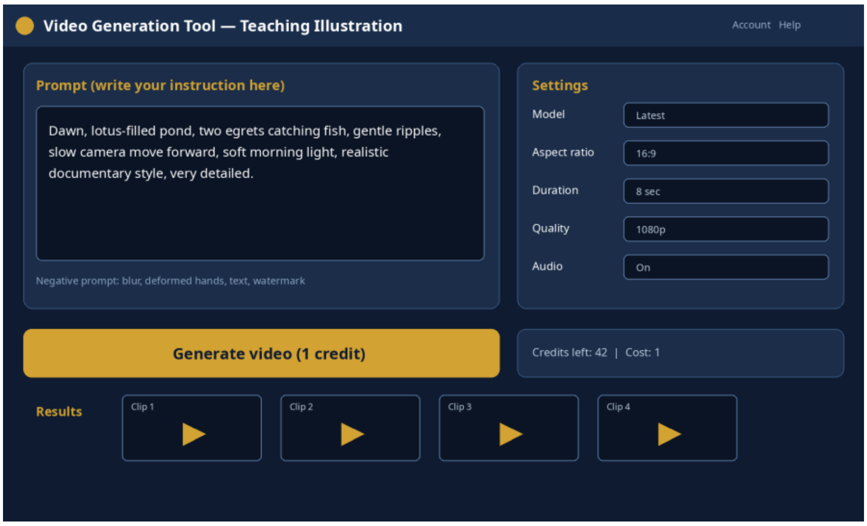


Figure 5 — Typical Layout of a Text-to-Video Tool (Teaching Illustration): Prompt Area, Settings, and Result

Now it is time to work with an actual tool. In this chapter, we will understand the common structure of text-to-video tools and create a first clip.

First, understand the landscape. This field changes extremely quickly: new models appear every few months and older ones become more capable. At the time of writing (2026), prominent names include Google Veo, OpenAI Sora, Kling, Runway, PixVerse, Seedance, Hailuo, Luma, and Pika. Each has a somewhat different character — some are strong at realistic scenes, some at stylized or artistic imagery, some are faster or less expensive. The names will continue to change, but the underlying structure described below is similar across tools. Learn the structure, not the brand names.

Look at Figure 5. Nearly every tool contains these five elements:

One — prompt area: a large empty box where you write your instruction. This is the heart of the tool.

Two — model selection: one company may offer several models — a newer and more powerful one that costs more, and an older or “fast” model that costs less. While learning, use the cheaper model; use the better model for final, publication-quality clips.

Three — aspect ratio: 16:9 for YouTube and television (wide); 9:16 for Reels, Shorts, and status videos (vertical); 1:1 for square output. Decide the destination at the beginning, because changing aspect ratio later can crop the scene.

Four — duration: often 5, 8, or 10 seconds. Short clips use fewer credits and tend to contain fewer errors.

Five — the Generate button and result area: press the button, wait from a few seconds to a few minutes, and the clip appears in the result area. Download it from there.

Now make your first clip, step by step:

1. Open the tool’s website and create an account with your email address.

2. Check the free credits — how many you have and when they renew.

3. Choose the cheaper/fast model, 16:9 aspect ratio, and the shortest available duration.

4. Paste the prompt you prepared in Chapter 5.

5. Press Generate and wait.

6. Watch the complete clip — not once, but two or three times. Pay attention to hands, faces, lettering, and movement.

7. Download it and save it in your “Clips” folder with a meaningful name — even if imperfect, it may be useful for comparison.

8. Run the refinement cycle: identify the weakness, improve the prompt, and generate again.

A few practical tips: generate two or three clips from the same prompt because AI produces a slightly different result each time, and you can select the best. If the main subject is correct but there is an error near the edge of the frame, the defect may be cropped out in editing — do not discard the clip immediately. And before you begin spending credits, set a daily experimentation limit; enthusiasm can consume a week’s credits in a single evening.

Exercise: generate the clip for the first scene of your storyboard. Run at least three refinement cycles. Save the final clip together with the exact prompt that produced it.

Image-to-Video: The Route to Greater Control

Text-to-video has one disadvantage: you do not fully control the starting image. The AI decides what the first scene will look like. Image-to-video solves this problem: first prepare a still image that matches your intention, then ask AI to animate it. The visual design, character’s face, and clothing are already fixed; AI mainly adds motion.

Where can the image come from? There are three main sources:

First — your own photograph. Animating your own photographs of a village, festival, nature, or art is the most authentic route. Remember that the photograph should either be your own or used with the photographer’s permission; and before animating the photograph of a living person, obtain that person’s consent — this is both good practice and part of the ethical principles discussed in Chapter 15.

Second — an AI-generated image. There are dedicated image-generation tools, and many video tools also include image generation. You can use the same six-element prompt logic, replacing camera movement with composition. Images are faster and cheaper to generate, so refine the still image first: creating ten images and selecting the best one is far cheaper than creating ten videos.

Third — a scan of your own artwork. A Mithila painting or other work that you created yourself or have the right to use can be scanned and given subtle motion — for example, a fish may move gently or a line ornament may shimmer. Use great restraint here. The dignity of traditional art often lies in its stillness, so keep movement extremely subtle and explicitly use phrases such as “gentle, slow, subtle motion” in the prompt.

The method is simple. Select the “image-to-video” option, upload your image, and write a motion instruction. Here, describe the existing scene less and the motion more: what should move, in what direction, how quickly, and what the camera should do. For example: “gentle ripples on the water, the egret’s wings moving slowly, the camera making an extremely slow push-in, nothing else changes.” The final phrase — “nothing else changes” — is important; without it, AI may sometimes alter the entire scene.

This is one of the best methods for solving character-consistency problems. If the same character appears repeatedly, first create or select one excellent image of that character, then animate that

same image with different motion instructions for each scene. Some advanced tools offer features called “character reference” or “consistent character,” where you provide the character image once and the system tries to preserve that character across clips. Look for such a feature in the tool you use.

Exercise: take one image — your own photograph or an AI-generated image — and make three motion versions: one where only the camera moves, one where only an element within the scene moves, and one where both move. Compare all three. Which feels most natural?

Voice: Voice-Over and Text-to-Speech

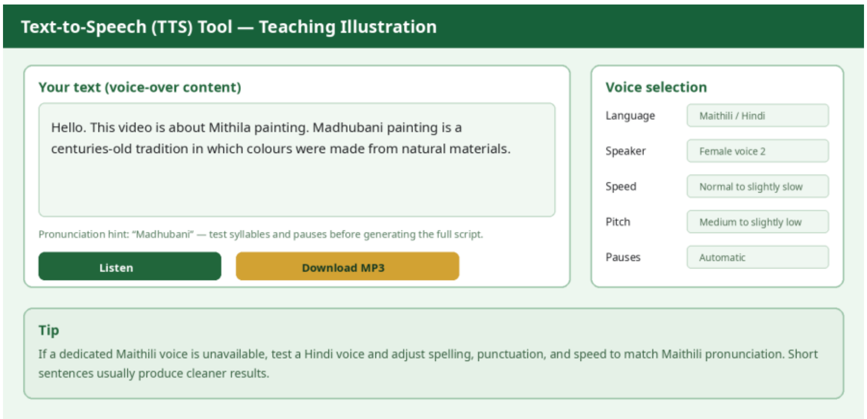


Figure 6 — Layout of a Text-to-Speech (TTS) Tool (Teaching Illustration)

The soul of a video often lies more in its voice than in its images. Viewers may tolerate a slightly imperfect visual, but they will stop watching a video with poor sound. Read this chapter carefully — and for Maithili speakers, several special points are especially important.

There are two ways to add a voice: record your own, or generate one with AI using text-to-speech (TTS).

Let us begin with your own voice, because for Maithili it is still often the best route. No machine can yet reproduce correct Maithili pronunciation, natural accent, and emotional modulation as reliably as a native speaker. You do not need a studio: the microphone in a modern smartphone is often good enough. Follow a few rules: record in a quiet room, with the fan and windows closed, preferably at night or early morning; keep the phone roughly a handspan from your mouth; stand or sit upright so the voice remains open; keep the script in front of you but speak as if telling, not reading; and record paragraph by paragraph rather than the whole script at once, so that a mistake requires only a small retake.

AI can improve a recorded voice. Many tools — often called “Enhance Voice” or “Studio Sound” inside editing apps — can remove background noise and make an ordinary recording sound more studio-like. Use this capability when appropriate; the difference between raw and enhanced speech can be substantial.

Now consider text-to-speech, as shown in Figure 6. You enter the text, choose a language and speaker voice, adjust speed or tone, and download an MP3 file. For Hindi, English, and many major languages, this technology is highly mature. Support for Maithili is improving; some tools are beginning to offer Maithili voices, and services built for Indian languages may expand faster in this direction. Check the tool you use. If Maithili is available, first test a short passage to see how accurately it pronounces the language.

If a Maithili voice is not available, there are two options. The first and best is to use your own recording as described above. The second is to use a Hindi voice as a compromise: write the text phonetically, listen closely and adjust spellings until the pronunciation improves, keep the speed somewhat slower, and use short sentences. The result will not have a fully natural Maithili accent, but it may be serviceable. Remember that this is a compromise, not an ideal solution; wherever emotion and linguistic accuracy matter — poetry, stories, children's material — use your own voice.

A newer capability is voice cloning. Some tools can build a digital representation of your voice from a few minutes of recording and then speak any text in that voice. This can be attractive for Maithili content because one trained voice can support many videos. But follow two strict rules: clone only your own voice — imitating anyone else without written permission is unacceptable — and where a cloned voice is used, disclosure is good practice.

The voice and visuals will finally be synchronized during editing in Chapter 11, so keep the voice file separately in the "Voice-Music" folder. Create the voice-over before generating all the video clips; once you know the narration length, you can judge how many seconds each scene needs.

Exercise: create the voice-over for your script in both ways — once by recording your own voice and once using a TTS tool. Listen to both with your eyes closed. Which one feels more like "you"? Why?

Avatar Video: The Digital Presenter

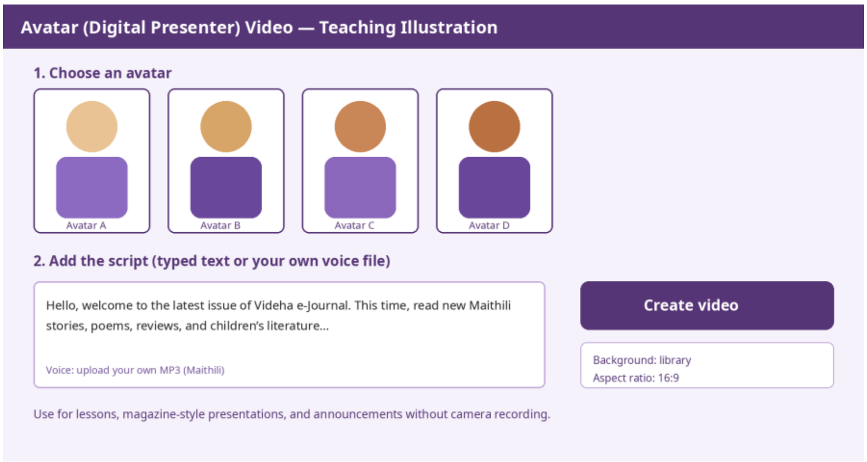


Figure 7 — Layout of an Avatar Video Tool (Teaching Illustration): Avatar Selection and Script

Imagine that every fifteen days you need to make an announcement video for a new issue of a magazine, or fifty lectures for a course. Setting up camera, lighting, clothing, and recording every time may be impractical. Avatar video offers one solution: a digital human appears on screen and speaks your written text — moving lips, blinking, and gesturing — much like a newsreader.

Well-known tools in this category include HeyGen and Synthesia, and the category is growing rapidly. Their basic structure is broadly similar, as shown in Figure 7:

One — choose an avatar. Tools may offer hundreds of ready-made avatars of different ages, appearances, and poses. Some tools also let you create an avatar from your own photograph or short video — allowing "you" to appear on screen without recording every time. If you create a personal avatar, apply the same consent rule used for voice cloning: use your own likeness, not someone else's without permission.

Two — provide the script. This may be written text that the tool speaks through TTS, or your own recorded voice file to which the avatar lip-syncs. For Maithili, the second route is often better: upload your own Maithili voice-over and let the avatar match its lip movements. The result can be surprisingly effective.

Three — choose the background and layout: library, office, plain colour, or your own image. Select the aspect ratio, such as 16:9 or 9:16, and generate.

The strength of avatar video lies in its practicality, not in spectacle. It is excellent for news-style presentation, magazine announcements, lesson explanations, institutional introductions, and information videos. For emotionally driven storytelling and poetry, a human presenter is still usually better.

One presentation tip: do not keep the avatar continuously on screen for the entire video. From time to time, show related visuals, photographs, or text cards — known as B-roll — while the avatar's voice continues underneath. This makes the video feel more alive. The combination happens in editing, so create and keep the avatar clips and B-roll clips separately.

And yes — when an avatar looks like a real human, viewers have a right to know that it is a digital presenter. Add one line to the description. This preserves trust, and trust is the real capital of a channel.

Exercise: using the free plan of an avatar tool, make a 30-second introduction video on "My Village" or a favourite book. Try both methods: typed text and an uploaded voice recording.

Music and Sound Effects

Visuals are for the eyes, speech for the ears, and music for the heart. The same scene can feel empty without music and come alive with the right score. But music also brings one of the greatest risks: copyright. If you use someone else's recorded song without permission, YouTube may block the video, divert revenue to someone else, or penalize the channel. So remember one rule: never use a well-known film song or commercial recording unless you have explicit permission.

Where, then, can music come from? There are three lawful sources:

First — AI-generated music. There are tools that create music from a written description: write "slow, emotional, flute-led, 60 seconds," and a track is generated. Some tools can also create a complete song, including vocals, from lyrics. Music created for your own video may often be usable, but read the licence terms of the tool at least once — especially whether commercial use, including monetized YouTube publication, is allowed.

Second — copyright-free or openly licensed music libraries. YouTube's own Audio Library inside YouTube Studio is free and generally safe for platform use. Many other services provide openly licensed music; some are entirely free, while others require attribution. If attribution is required, do not forget to credit the musician in the video description.

Third — music you record yourself. Mithila has a rich musical tradition. If you or someone you know sings or plays, your own recording of a traditional melody can be the most authentic source and can give the video an identity that no AI can provide. Remember the distinction: a folk melody may be traditional, but a particular person's recording or arrangement is their own work.

The art of using music: keep background music soft beneath voice-over — roughly twenty to thirty percent of the perceived level of the main speech. Where there is no voice-over, such as an

opening, ending, or visual transition, the music may rise. Match the emotional character of the music to the scene — freshness at dawn, pathos at farewell. At the end, let the music gradually fade out; an abrupt stop can feel harsh to the ear.

Sound effects are small environmental sounds — birdsong, a splash of water, the rustle of wind. They make a scene more believable. Some newer video models generate sound together with the visuals as "native audio." If your tool offers this and it suits the project, you may use it; otherwise, obtain rights-cleared sounds from a sound library and add them during editing.

Exercise: find two different styles of background music for your video — one AI-generated and one from a rights-cleared library. Imagine or test each beneath the video. Which better matches the intended mood?

Editing: Turning Clips into a Video

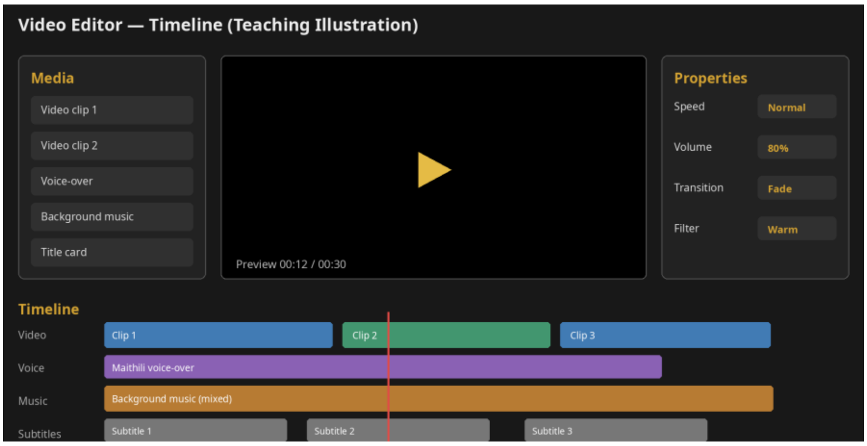


Figure 8 — Editor Timeline (Teaching Illustration): Tracks for Video, Voice, Music, and Subtitles

Now you have all the ingredients — video clips, voice-over, and music. Editing is the kitchen where these ingredients become a finished dish. The good news is that editing skills, once learned, remain useful across projects; they do not change as rapidly as AI tools.

Choose an editor. Free editors are available for both mobile devices and computers. CapCut is currently one of the most common and accessible options; on computers, DaVinci Resolve offers professional-level capability even in its free version. Choose whichever suits you — the basic structure shown in Figure 8 is similar across editors.

Media area: where you import your files. Preview: where the video plays. Timeline: the most important area — a line of time on which clips are arranged in order. The timeline contains several tracks layered above one another: video, voice, music, subtitles, and so on, all playing together.

Use this basic editing order:

1. Place the voice-over on the timeline first. It is the spine of the video; the visuals will be arranged around it.

2. Listen to the voice and place the video clips in sequence so that the picture corresponds to what is being said. Trim each clip and keep only its strongest section. AI clips often contain awkwardness at the beginning or end; trimming both edges can greatly improve them.

3. Add transitions between clips. Rule: the fewer, the better. A direct cut is the cleanest; a soft fade can mark a shift in mood. Spinning, bouncing, brightly coloured transitions often make a video look amateurish.

4. Place the music on its track and lower its level beneath the narration, as discussed in Chapter 10.

5. Match the colour. Most editors have a one-click filter or enhancement option. If clips come from different AI tools, applying the same restrained treatment can bring them closer in colour and make the video feel more unified.

6. Add a title card of 3–4 seconds at the beginning and an end card with the channel name or a simple “Thank you” at the end.

Export settings: 1080p, MP4 format, and 30 frames per second are common settings for YouTube. Save the exported file in the “Final” folder.

One final editing habit: watch the whole video from beginning to end without stopping, as an ordinary viewer. Wherever your attention wanders, the sequence probably needs trimming or improvement. Then show it to someone at home. The first reaction of a real viewer may teach more than many abstract critiques.

Exercise: combine all your clips, voice, and music into your first complete video, including title and end cards. Export it, show it to a family member, and write down the person’s first reaction.

Subtitles: Making Video Readable

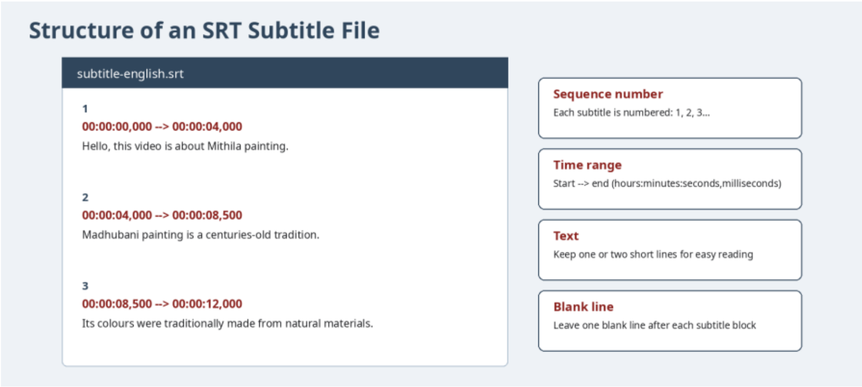


Figure 9 — Structure of an SRT Subtitle File

Today, many viewers watch video without sound — on a bus, at work, or in bed at night. Without subtitles, you may lose them. For Maithili content, subtitles matter twice over: Maithili subtitles strengthen reading practice in the language, while Hindi or English subtitles bring in viewers who do not understand Maithili. In this way, stories from Mithila can reach audiences far beyond the language community.

The standard subtitle format is SRT — a simple text file in which three things repeat: sequence number, time range, and subtitle text, as shown in Figure 9. The file can be created or corrected even in an ordinary text editor such as Notepad.

Matching every subtitle to time by hand is labour-intensive, and AI can help again. There are two routes:

First — automatic transcription. Editors such as CapCut offer “auto captions,” which listen to the video and create timed subtitles automatically. For Hindi and English, accuracy can be high. A Maithili voice may sometimes be interpreted as Hindi, so you will need to correct the generated text. Even so, correcting an already timed transcript is much faster than writing and timing everything from scratch.

Second — from the script. You already have a written script. Give the script and total video duration to a chatbot and ask: “Divide this into SRT format, with no subtitle longer than two lines and with comfortable reading speed.” Then import or recreate the timings in the editor and adjust them forward or backward as needed.

Craft rules for subtitles: use no more than two lines at once, with roughly thirty-two to forty characters per line. Keep text on screen for at least about one second. Break lines at meaningful phrase boundaries — not “I went / to the market,” but “I went to the market” together where possible. Use light-coloured text with a subtle dark shadow or background strip so it remains readable over bright images. Place subtitles near the lower centre, but not so low that they are cropped or covered in vertical-video layouts.

On YouTube, subtitles can be provided in two ways: burned into the video during export, or uploaded as a separate SRT file that viewers can turn on or off. A strong method is to upload Maithili as a separate subtitle file, then add separate Hindi and English SRT versions as well. A chatbot can assist with translation, but you are responsible for checking the translation. One video can then reach audiences in three languages.

Exercise: create a Maithili SRT file for your video using either method, then create a Hindi translation file. Play both with the video and check the timing and line length.

Publishing on YouTube

YouTube Upload Checklist

Title	Clear Maithili title + short English keyword when useful; keep it concise.
Description	Put a two-line summary first, then sources, links, and channel information.
Tags	Use relevant terms such as Maithili, Mithila, and subject keywords in English and Devanagari as needed.
Thumbnail	1280×720; large readable text; no more than a few words; show the face or main subject clearly.
Language	Choose Maithili as the video language when the platform supports it; upload subtitle files.
AI disclosure	If visuals are AI-generated or materially altered, add an appropriate synthetic/altered-content disclosure.
Playlist	Add the video to a topic-based playlist so viewers can easily find the next related video.

Figure 10 — YouTube Upload Checklist

The video is finished; now it must reach viewers. YouTube remains one of the largest and most durable homes for video: a video posted there can continue to be watched for years, while content on short-form platforms may have a much shorter life. Treat YouTube as the main home and Reels or Shorts as its windows.

Creating a channel is simple: sign in to YouTube with your Google account and create a channel. Choose the channel name carefully — short, easy to pronounce, and suggestive of the subject. A channel logo and banner can also be created with appropriate design or image tools.

Keep the checklist in Figure 10 beside you while uploading. A few points deserve expansion:

Title: put the main subject in the first three or four words. Adding a Hindi or English version in brackets beside a Maithili title may improve discoverability because viewers search in multiple languages.

Description: the first two lines are especially valuable because they often appear in search previews. Put a concise summary of the video there, including important keywords. Below that, add chapter timings, sources, and a short channel description.

Thumbnail: viewers often see the thumbnail before they read the title. Use one image, one strong idea or emotion, and no more than about three words — large enough to read on a small phone screen. You may use AI image tools for an attractive thumbnail, but do not mislead: what appears in the thumbnail should actually be represented in the video, otherwise viewers feel deceived and trust declines.

AI disclosure: when a platform asks whether realistic-looking content has been synthetically created or altered, answer transparently. This is not merely rule compliance; it is honesty with the audience. Clearly fictional animation may be different, but when in doubt, disclosure is the safer practice.

Language setting: select "Maithili" as the video language when the platform supports it. This helps the video reach people searching for Maithili and strengthens the visibility and data footprint of Maithili content.

What after uploading? Observe the response during the first hours and days and reply to comments; early conversation helps build the roots of a channel. Maintain regularity as well. One video every fortnight means twenty-four videos a year, and a regular series can be more valuable than a hundred scattered uploads. Audiences connect with a programme, not randomness.

For Reels or Shorts, cut the most engaging 30–60 seconds of the main video into a 9:16 version using reframe or crop tools in the editor. At the end, add a simple message such as "Full video on the channel." Short-form content can become the window that brings new viewers to the longer work.

Exercise: upload your video after completing all seven points of the checklist. Also cut and upload a separate Shorts version. After one week, compare the basic figures, such as views and watch time, for both.

Special Considerations for Maithili Content

This chapter is the heart of the book. The tools are universal, but our purpose is particular: to create in Maithili and for Mithila.

First — linguistic accuracy. AI tools often introduce Hindi influence when writing Maithili because they have typically seen far more Hindi. Scripts, translations, and subtitles produced by chatbots must therefore be checked by you. A simple rule: AI-generated Maithili is a draft, not an authority. When in doubt, trust your ear. Say the sentence aloud; if it jars, examine whether Hindi syntax or vocabulary has slipped in.

Second — the question of script. Maithili is widely written in Devanagari and also has its own historical script, Tirhuta or Mithilakshar. For subtitles and general on-screen text, Devanagari will usually reach the largest audience. Use Tirhuta for identity and ornament — on title cards, logos, or watermarks — so that viewers continue to encounter the script and curiosity about it grows. Remember that AI image generators may not reproduce Devanagari or Tirhuta lettering accurately; add important text yourself during editing rather than asking the image model to write it.

Third — visual authenticity. If you prompt AI with only "Indian village," it may create a generic North Indian scene that is not specifically Mithila. Include concrete Mithila markers in the prompt: pond, banyan tree, mango grove, paddy fields, fish, betel, thatched house, tulsi altar in the courtyard, aripan-like wall ornament, and other contextually appropriate details. Even then, the result may be "Mithila-like," not Mithila itself. Whenever possible, mix in genuine photographs and video. A combination of AI visuals and real footage often produces the most authentic result.

Fourth — respect for Madhubani/Mithila painting. An AI can be instructed to create an image "in Madhubani style," and it may imitate some colour and line patterns. But pause and consider the living tradition. Mithila painting supports the livelihoods of thousands of artists, and forms such as Bharni, Kachni, Godna, and Gobar have their own histories and lineages. An AI-made "Madhubani-like" image borrows the surface appearance without the labour and cultural knowledge embedded in the tradition. My recommendation is to show the real work of real artists, with permission and credit. This both respects the artist and strengthens the video. If you use AI-generated Madhubani-like ornament, say clearly that it is an AI-made imitation, not an authentic Madhubani work.

Fifth — an almost inexhaustible store of subjects. The field of Maithili video remains comparatively open, and many creators can still be among the first to build sustained series. Possible directions include children's content — songs, stories, literacy; explanations of festivals such as Chhath, Sama-Chakeba, and Jur Sital; recipes; illustrated presentations of folktales or Vidyapati songs; a visual dictionary of village and household words that are disappearing; and introductions to places in Mithila. Each direction could support a dedicated channel.

And finally — patience with quality and scale. A Maithili channel will naturally have a smaller potential audience than a Hindi channel, but Maithili viewers may form deeper relationships with creators: commenting, sharing, and returning. Do not look only at numbers; look at relationships. One hundred committed viewers can be more valuable than ten thousand indifferent ones.

Exercise: choose one of the directions above and plan three consecutive videos on it, each with a topic and one-line summary. One video is an experiment; three begin to reveal a direction.

Copyright, Ethics, and Responsibility

Powerful tools bring responsibility. This chapter is short, but put every line into practice.

Copyright — the basic principle: do not use what you did not create without permission. This applies to film songs, other people's video clips, book text, photographs, and recordings. "Everyone does it" is not a defence. Automated systems such as YouTube Content ID can identify protected material and the channel may bear the consequences. Even with openly licensed material, read the conditions: some licences require credit, while others prohibit commercial use.

Understand the rights attached to AI-generated material as well. Tool terms may differ regarding commercial use of generated video or images, and conditions can vary by plan. If you use a tool regularly, read its terms of use at least once, especially sections dealing with ownership and commercial use.

Respect for persons — three strict prohibitions: do not imitate someone's face or voice without written consent, whether the person is living or dead; do not place a person in a scene that degrades them or put words in their mouth that they never said, as in a deepfake; and exercise

special care with images and voices of children. Even before publishing a video of your own child, remember that publication makes it public.

Commitment to truth — AI can create realistic-looking scenes, and therefore also create confusion. Keep fiction clearly fictional and news clearly factual. For historical or cultural subjects, check facts against reliable sources. AI chatbots can make confident errors, commonly called hallucinations. Where an AI-generated scene resembles reality, disclose it appropriately. Viewer trust may take years to earn and one misleading video to lose.

And responsibility to yourself — never let convenience become intellectual laziness. Use AI for work, not for thinking in your place. The day you hand over script, selection, and final judgment entirely to AI, the video will no longer be fully yours — and viewers can sense the difference. Keep the tool in your hand; do not become an extension of the tool.

Exercise: conduct a “responsibility check” on your first video. Do you have the music licence? Permission for image sources? Fact checking? AI disclosure? Keep written answers to all four. Repeat this check for every video until it becomes a habit.

Practice Project: Your First Video from Beginning to End

Now bring all the principles together. This chapter presents a complete example project: a 90-second video titled “Chhath: The Great Festival of Sun Worship.” Read the process, then repeat the same sequence with your own subject.

Day 1 — planning (Chapters 3–4): objective — explain the basic spirit and practice of Chhath simply. Audience — Maithili families living away from Mithila, especially the younger generation. Duration — 90 seconds, meaning roughly nine or ten clips. A chatbot was used to draft a two-column script, then three parts were manually revised: local detail was added to the description of “Nahay-Khay,” and the final line was given greater emotional depth. The storyboard contained ten scenes: a river ghat at dawn for establishment, a courtyard where thekua is being made, baskets and offerings for arghya, evening arghya, morning arghya, paran, and a closing title card.

Day 2 — voice (Chapter 8): the voice-over was recorded in the creator’s own voice at night in a quiet room, paragraph by paragraph. The editor’s “Enhance Voice” feature was used to clean it. The result was 82 seconds long, confirming that the scene plan would fit.

Days 3–4 — visuals (Chapters 5–7): a six-element prompt was prepared for every scene, conceived in Maithili and translated into English. Each prompt included specific Mithila markers such as “river ghat in North Bihar, banana-stem decoration, bamboo basket, fruit in a soop-daura.” Three clips worked well on the first attempt; four clips needed two or three refinement cycles. For two scenes — thekua and the offering basket — the image-to-video method was used: first improve the still image, then add subtle motion. One scene used an actual photograph of the crowd at the ghat because AI could not reproduce that lived density convincingly. This mixture of synthetic and real material is one of the strongest approaches discussed in the book.

Day 5 — music and assembly (Chapters 10–11): AI music was generated with the description “slow, devotional, flute and mridang, 90 seconds,” and the licence was checked. The edit was built in this order: voice first, then clips, mostly direct cuts with two fades, and one restrained warm colour treatment applied across the clips. The title card used “Chhath” in Tirhuta and the full title in Devanagari.

Day 6 — subtitles and publication (Chapters 12–13): a Maithili SRT was made from the script, followed by Hindi and English translation files, with chatbot assistance and human checking. The thumbnail used an evening arghya image with three words meaning "Why Chhath?" The upload language was set to Maithili, AI disclosure was enabled where appropriate, and music credit and sources were added to the description. A 35-second Shorts version was also prepared.

Day 7 — review (Chapter 15): the responsibility check was completed on all four points. The family viewed the video and suggested that the morning arghya scene felt too short; that observation was recorded for the next video.

Notice the scale: seven days, one or two hours per day, from zero to a published video. The first project may take longer — allow it. By the third or fourth video, the sequence begins to become habit and the time required may fall sharply.

One final point. The book ends here, but learning begins here. Tools will change, models will change, screens will change, but the structure you have learned — planning, prompting, refinement cycles, assembly, and responsibility — will remain useful. Now go and tell the stories of Mithila to the world, in your own voice and your own language.

Glossary

AI (Artificial Intelligence) — the capacity of a computer to learn, understand, and perform tasks in ways associated with human intelligence.

Generative AI — AI that creates new content such as text, images, video, or voice.

Prompt — a written instruction given to an AI system.

Negative prompt — a list of things that should not appear in the result.

Text-to-video — a method of creating video from a written description.

Image-to-video — a method of animating a still image.

TTS (Text-to-Speech) — a method of turning written text into spoken audio.

Avatar — a digital presenter that speaks text on screen.

Lip-sync — matching mouth movement to spoken audio.

Voice cloning — creating a digital imitation or model of a person's voice.

Credit — a unit of usage in an AI tool; a certain number may be consumed for each generation.

Model — a particular version or engine of an AI system.

Clip — a short section of video, often about 5–20 seconds.

Aspect ratio — the relationship between screen width and height; for example, 16:9 wide or 9:16 vertical.

Resolution — the level of image detail; 1080p is a common standard, while 4K provides greater detail.

Watermark — an identifying mark placed on media.

Storyboard — a scene-by-scene visual plan.

B-roll — supporting visuals shown in addition to the main presenter or interview.

Timeline — the time-based area in an editor where clips are arranged.

Track — a layer on the timeline for video, voice, music, and other elements.

Cut — trimming a clip or moving directly from one clip to another.

Transition — the method or effect used to move between two clips, such as a fade.

Fade in/out — gradually appearing from or disappearing into darkness or silence.

Render/export — converting an edited project into a final video file.

SRT — a standard subtitle file format.

Burned-in — subtitles permanently embedded in the video image.

Thumbnail — the cover image used to represent a video.

Hallucination — a confident but false output produced by AI.

Content ID — YouTube's automated copyright-identification system.

Deepfake — a deceptive synthetic imitation of a person's face or voice.

Monetization — earning income from video content.

Prompt Collection

Below are ten ready-to-use prompt templates, presented here in English translation. Fill the bracketed parts with your own details and adapt them to the method explained in Chapter 5.

1. Village Dawn: "A village in North Bihar at dawn, light mist over the fields, thatched houses, a banyan tree in the distance, [your detail], camera slowly panning right, soft golden light, realistic documentary style."

2. Pond Scene: "A pond filled with lotus and makhana leaves, [season], [tree/ghat] on the bank, gentle ripples in the water, subtle movement of [bird/fish], camera moving forward extremely slowly, [time of day] lighting, cinematic."

3. Courtyard Scene: "A courtyard of a Mithila home, tulsi altar, [action — for example, food cooking on a hearth], folk ornament on the wall, warm intimate light, close-up, realistic."

4. Festival Scene: "Preparations for [festival name], [main object — for example, soop-daura, lamps], courtyard/ghat setting, close-ups of busy hands, festive joy, warm lighting, documentary style."

5. Nature Scene: "Green paddy fields moving in the wind, [time], distant horizon, [bird] flying, drone shot slowly rising, natural light, highly detailed."

6. Object Introduction (Food/Craft): "Extreme close-up of [object], [background], subtle steam/glow, camera slowly circling the object, soft studio-like light, appetizing/appealing style."

7. Story Illustration (Animation): "[Character description] [performing action], [location], hand-drawn animation style, soft watercolor, children's-book illustration feel, slow pleasant motion."

8. Historical Imagination: "An imagined scene of Mithila in [period], [action], colours inspired by old palm-leaf illustrations, slow serious camera, museum-documentary style." Add a clear AI disclosure that this is an imaginative reconstruction.

9. Title-Card Background: "Abstract ornamental pattern moving very slowly, [colour scheme — for example, vermilion red and yellow], no letters, no figures, calm uniform movement, 10-second loop." Add the title later during editing.

10. Motion Instruction for Image-to-Video: "In this image, add gentle natural motion only to [element — for example, water/leaves/fabric], camera [static/extremely slow push-in], change nothing else, preserve the original colours and linework."

Tool Guide (2026 Context)

This list reflects the situation at the time of writing in 2026. Names and features change, so learn the category rather than memorizing brands. Before adopting a tool, check three things: what the free plan offers, whether commercial use of generated material is permitted, and the watermark policy.

Text-to-video / image-to-video: Google Veo, OpenAI Sora, Kling, Runway, PixVerse, Seedance, Hailuo, Luma, Pika — tools for realistic scenes and cinematic clips.

Avatar video: HeyGen, Synthesia — digital presenters, lip-sync, and multilingual output.

Text-to-speech and voice cloning: ElevenLabs and related services; Indian-language-focused services are also expanding, so continue watching for improved Maithili support.

AI music: Suno, Udio, and similar tools — music or songs generated from descriptions or lyrics.

Image generation: included in many video tools; separate tools include Midjourney and image features integrated with chatbots.

Editing: CapCut — mobile and computer, accessible, with auto captions; DaVinci Resolve — computer-based, free tier, professional-level; Canva — thumbnails, banners, and simple video.

All-in-one planning-to-video tools: systems such as InVideo can combine script, visuals, and voice from a topic for rapid production, but may reduce your direct creative imprint. For learning, the step-by-step method in this book remains preferable.

Chatbots for scripts, translation, SRT, and brainstorming: Claude, ChatGPT, Gemini.

And finally — this book too will age, but you need not. The "Teach Yourself" habit developed here can be applied to every new tool. When you encounter a new system, ask three questions: where do I write the prompt? what settings are available? where and how does the result appear? With those questions, you can usually begin learning the tool within minutes.

Part 3
Advanced Technical Study

Artificial intelligence (AI) has brought a major transformation to the technological world, especially in video production. Where high-quality video once required expensive cameras, professional editors, motion designers, and hours of post-production, by 2026 many parts of the process can be completed in minutes. Creators, agencies, and businesses can now use a simple text prompt to generate professional-looking video, edit scenes, create AI avatars, and produce voice-overs. As part of the Teach Yourself series, this book offers a detailed study of the complete process of creating Maithili videos with AI, including tools, prompt engineering, voice cloning, and evaluation. It includes practical examples and explanatory interface screenshots to give readers a clear understanding of how AI video tools are used.

Leading AI Video Generator Tools in 2026

Choosing an AI video-generation platform can be complicated because each tool has its own strengths. Some specialize in realistic avatars, while others are stronger in cinematic scenes or advanced editing. No single tool is perfect, but excellent results can be achieved by combining different tools appropriately.

Google Veo and Its Advanced Versions

Google Veo is a highly capable platform known for cinematic AI video and high-quality visual storytelling. Developed by Google DeepMind, this video-generation model can create new video from text, image, or video references.

After its initial announcement in May 2024, Google released Veo 2, Veo 3, and later Veo 3.1. In the source text, Veo 3.1 is described as mobile-optimized, with vertical-video integration, higher resolutions up to 4K, and support for uploading multiple reference images, up to three. It also describes scene extension and controls for specifying first and last frames, allowing a short clip to be extended into a longer sequence.

Figure Description 1: the Google Veo 3.1 dashboard, showing an “Upload Reference Images” option and a drop-down menu for selecting aspect ratio.

The table below summarizes the technical development across these versions:

Feature	Veo 2	Veo 3	Veo 3.1
Audio	None; silent video only	Yes, available	Yes; richer soundscape and synchronization
References	No	No	Yes; up to three reference images
Video duration	Five to eight seconds	Eight seconds	Four, six, and eight seconds
Resolution	720p	720p and 1080p	720p, 1080p, 4K
Control features	Less precise prompt	Improved prompt	First- and last-frame

	adherence; limited lighting variation	adherence and camera motion	transitions and scene control
Quality	Good	Better; improved physics and camera motion	Excellent; continuity tools and audio integration

Plans, credits, resolution limits, watermark policies, and prices for AI video tools change frequently. While learning, begin with short tests, choose one or two tools according to your needs, and focus more strongly on durable skills — prompting, storyboarding, camera language, sound, and editing.

Runway Gen 4.5 and Creative Control

Runway is a powerful tool for filmmakers and VFX artists. It is designed for creators who are not satisfied with simply writing a prompt and instead want greater control over the creative process. The source text describes Runway Gen 4.5 as offering advanced camera controls such as pan, tilt, and zoom, together with features such as a multi-motion brush that can animate selected areas of a still image.

Its interface may feel somewhat complex to a new user. The sidebar contains tools, apps, chat, workflows, live features, and other options, so finding the prompt area can initially be difficult. The source text also notes that while background generation can be strong, character movement may sometimes look unnatural and glitches may appear around faces or eyes.

Figure Description 2: the Runway Gen 4.5 editing window, where a “multi-motion brush” is being used to mark a specific part of an image, such as river water, so that motion is generated only in that area.

The Speed of Seedance 2.5

Seedance, developed by ByteDance, is described in the source as well known in the AI-video community for speed. The source states that Seedance 2.0 could generate a 10-second clip in about 30 seconds, while version 2.5, referenced there through MuAPI, is more advanced. It is described as supporting video durations from 4 to 30 seconds and accepting inputs ranging from JPEG, PNG, and WebP to MP4 and MOV.

Seedance 2.5 is described as supporting a broad range of aspect ratios, including 16:9, 9:16, 1:1, and 21:9, making it useful for both social-media and cinematic video. For output, the source describes MP4 with H.264/yuv420p as the default, while MOV/yuv444p is offered for higher colour fidelity and repeated editing.

Other Major Video-Generation Platforms

Other platforms in the market also have distinctive uses and identities.

HeyGen: the source describes this tool as functioning like a complete production studio. It creates presenter-led video in which avatars can speak with accurate lip-sync across more than 175 languages.

VidCap and DeeVid AI: the source presents VidCap as especially useful for rapid, accurate captions and multilingual text overlays, while DeeVid AI is described as a fast option for turning images or text into short social-media videos.

LTX Studio: this tool is described as offering extensive creative control through shot-by-shot storyboarding, enabling scene-by-scene planning for an entire film.

KlingAI and Sora2: KlingAI is described as capable of producing highly realistic scenes, though occasional technical glitches may occur. OpenAI's Sora2 is described in the source as powerful for story-based clips, audio, and lip-sync, although output quality may at times be inconsistent.

The table below gives an overall comparison of leading AI video generators in 2026:

Category	Best For	Plan-related note
LTX Studio	Control through shot-by-shot storyboarding	Free/paid plans and pricing change over time.
Adobe Firefly	Commercially safer AI video workflows	Free/paid plans and pricing change over time.
Descript	Video editing through script editing	Free/paid plans and pricing change over time.
VEED	Rapid content production and repurposing	Free/paid plans and pricing change over time.
InVideo AI	Prompt-to-video for social media	Free/paid plans and pricing change over time.
OpusClip	Extracting viral-style clips from longer videos	Free/paid plans and pricing change over time.

Figure Description 3: the avatar-selection screen in HeyGen, displaying a list of realistic digital human avatars of different ages, genders, and clothing styles that can be selected with the mouse.

Principles and Science of Prompt Engineering

To obtain strong results from an AI video generator, it is necessary to understand prompt engineering. A good prompt is more like a descriptive picture than a bare command. Instead of writing only "make a beautiful sunset," a prompt such as "a vivid sunset over mountains, golden rays breaking through pink clouds, viewed from a low angle" gives the model much more useful visual information.

Formula for a Text-to-Video Prompt

A standardized formula can help prevent important details from being omitted: Prompt = Subject + Action + Setting + Camera + Lighting + Style + Format + Constraints.

The following explains how each element works:

Subject: this is the principal focus of the video. If the model cannot immediately identify the main subject, the output may become confused. The subject should be clear and specific, for example "a matte-black skincare bottle," "a young barista," or "a glass perfume bottle." Abstract phrases such as "a beautiful lifestyle," "a brand," or "a beautiful scene" should not be used as the main subject.

Action: video requires movement. A prompt without motion can feel like a still image. Give the subject one main visible action, for example: "the perfume bottle rotates slowly while soft smoke

rises behind it." Avoid giving five different actions to one short clip; if necessary, create separate clips and combine them in editing.

Setting: this gives the AI model contextual information for understanding the scene. It includes location, time of day, weather, background detail, and surface materials. Example: "inside a small independent bookshop at night, with wooden shelves and rain running down the window."

Camera: shot size, angle, and movement, such as close-up, slow push-in, or shallow depth of field.

Lighting: mood, time of day, and source of light, such as warm morning light or soft reflections.

Style: the visual form or genre, such as a realistic commercial product video or 3D animation.

Format: duration, aspect ratio, and delivery format, for example a 6-second vertical video for Instagram Reels.

Constraints: things that should not appear in the video. Negative prompts may include instructions such as "no text, no logo, no glitching."

Figure Description 4: the prompt box in InVideo AI, containing a detailed English prompt built from the formula above — Subject + Action + Setting + Camera — with a Generate button below it.

Multi-Shot Generation Formula

When telling a larger story, a multi-shot formula can be used, for example with models such as Wan 2.7/2.6. Multi-shot prompt = overall description + shot number + timestamp + shot content.

Begin with a short summary of the complete video. In each shot, use reference identifiers such as "Image 1" or "character1" so that a character's identity can remain consistent across clips. Dialogue lines should be written explicitly, for example X says: "xxx", so that the model can preserve the wording. If no speech is required, state "No dialogue" clearly.

Camera Angles, Movement, and Lighting

Communicating the language of cinematography to an AI video generator is an art in itself. Correct terminology can help transform a result from amateur-looking to much more professional.

Camera Angles

Eye-level shot: a neutral perspective, as though a person is viewing the subject from ordinary standing or seated height; for example, "an eye-level shot of a woman drinking tea."

Low-angle shot: the camera is below the subject and looks upward, making the subject appear powerful or imposing; for example, "a low-angle tracking shot of a superhero."

High-angle shot: the camera is above the subject, which can make the subject appear smaller, weaker, or part of a crowd.

Bird's-eye view: a shot taken directly from above, giving a map-like view of the whole scene.

Worm's-eye view: an extremely low shot looking directly upward from ground level, often used to emphasize the scale of skyscrapers or other towering objects.

Dutch angle: the camera is tilted to one side, creating a sense of unease, disorientation, or dynamic instability.

Camera Movements

Every camera movement should serve an emotional or narrative purpose.

Static shot: the camera remains completely still, suggesting calmness and focused attention.

Pan left/right: the camera rotates horizontally on its axis to the left or right. This is useful for introducing the viewer to a scene.

Tilt up/down: the camera rotates vertically on its axis, upward or downward.

Dolly in/out: the camera physically moves closer to or farther from the subject. A push-in can create intimacy or tension, while a pull-out can suggest separation or reveal scale.

Truck left/right: the camera moves horizontally, parallel to the subject, often alongside a moving character.

Pedestal up/down: the camera rises or lowers while maintaining its level, for example to reveal the height of a large tree.

Zoom in/out: the lens changes focal length while the camera itself remains in place. This differs from a dolly movement, where the camera physically moves.

Crane shot: a camera mounted on a crane is used for dramatic reveals or for showing large-scale settings such as a battlefield.

Whip pan: an extremely fast pan that produces motion blur, often used as a transition or to convey rapid movement.

Arc shot/orbit: the camera travels in a circular path around the subject, emphasizing the subject's importance or dimensionality.

Figure Description 5: examples of camera movement in which pan, tilt, and dolly are shown with directional arrows around a 3D model so the direction of movement can be understood clearly.

Lighting Parameters

Lighting is one of the most important factors determining video quality and mood. The source text emphasizes that the description of lighting can have a particularly strong effect on the result.

Cinematic warm light: use a phrase such as "golden-hour cinematography" to evoke the warm light of dawn or evening.

Chiaroscuro: deep shadow and high contrast between light and darkness for dramatic or mysterious scenes.

Three-point lighting: a balanced combination of key light, which establishes the principal direction and colour of light; fill light, which softens shadows; and backlight or rim light, which separates the subject from the background by outlining hair or shoulders.

Instead of writing only "dramatic lighting," it is more useful to identify a specific source, for example: "a warm key light from screen left across the avatar's face, with a soft pale backlight from behind separating the figure from the background."

Audio, Text-to-Speech, and Voice Cloning

A high-quality video is not complete without strong audio. AI audio technology has advanced rapidly, and platforms such as ElevenLabs are prominent examples in the source text.

Text-to-Speech Models and ElevenLabs

The source describes ElevenLabs as capable of producing text-to-speech in more than 70 languages, including Hindi and Indian regional or "Desi" accents. It notes that a directly pre-trained model explicitly labelled "Maithili" may not be universally available, but suggests that Hindi or Desi-accent voices, or a custom cloned voice, can be used for Maithili video. It further states that Maithili text written in Devanagari may be spoken reasonably clearly by a Hindi/Desi model.

The source lists several ElevenLabs models for different needs:

Flash v2.5: described as an ultra-low-latency model, around 75 ms, supporting 32 languages and up to 40,000 characters. It is presented as suitable for developer-oriented applications where speed is important.

Turbo v2.5: described as offering a strong balance between high quality and low latency, approximately 250–300 ms.

Eleven Multilingual v2: described as an especially realistic and emotionally expressive model for narration, audiobooks, and post-production across 29 languages, suitable for longer generations up to 10,000 characters.

Figure Description 6: the ElevenLabs Voice Studio interface, with Stability and Clarity sliders visible in Settings and some Devanagari text entered in the central text box.

Voice Cloning

Voice-cloning technology can create a digital replica of a person's voice. It may be used for podcasts, video voice-over, or audiobooks.

Instant Voice Cloning: the source describes this as capable of creating a voice replica within seconds from about 1–5 minutes of clear audio samples.

Professional Voice Cloning: the source describes this as requiring around 30 minutes or more of clean audio and as being able to capture fine vocal detail, pitch variation, and emotional rhythm. It also describes enterprise-grade security measures such as encryption and compliance frameworks including SOC 2 and HIPAA.

Emotional Control and Context

Modern AI voices do more than read text mechanically; they can respond to emotional cues embedded in the text. Adding directional tags may noticeably change delivery. For example:

"The voice paused for a moment, [softly] as if gathering its thoughts."

Tags such as [laughs warmly] or [sighs contentedly] can add natural laughter or breathing sounds to an AI voice and create a stronger sense of human presence for the listener.

The source states that Indian enterprises use such technology for customer service in sectors such as telecom and e-commerce, healthcare appointment scheduling, and multilingual audiobook generation, and mentions integrations with Indian telephony providers such as Plivo and Ozonetel.

Digital Avatars and Lip-Sync

Once both video and audio have been generated, the next important step is to combine them seamlessly, especially when a digital avatar or character is speaking on screen.

Lip-Sync Technology and Avatars

Accurate lip-sync is essential for a convincing digital avatar. The source lists tools such as HeyGen, Wav2Lip, and Sync Labs as systems that can match generated audio in different languages to the mouth movements of a video character. This is especially useful for localization into multiple languages.

Lens choice is also important for the visual language of a cinematic avatar:

A 35–50 mm lens look is commonly used for ordinary conversation and reactions.

A 24–35 mm lens is suitable for contextual shots in which more of the environment should be visible.

A 50–85 mm lens look is often used for beauty close-ups or product hand-offs.

Frame Rate and Clip Duration

A common misconception is that a higher frame rate such as 60 fps automatically makes video more cinematic or smoother. The source text notes that current text-to-video models are commonly oriented toward delivery in the 24–30 fps range.

Frame rate: for a cinematic feel and natural motion blur, the source recommends 24 fps. Requesting 60 fps may lead to interpolation and can introduce unnatural motion blur or artifacts. Use 25 or 30 fps when required by a social platform, broadcast workflow, or project specification.

Clip duration: the ideal length of a clip depends less on the platform than on the model's temporal stability.

A 2–3 second clip can be stable but may be too short to show meaningful emotional change.

A 4–6 second clip often provides a useful balance between narrative clarity and stability.

In clips longer than about 8 seconds, identity, lighting, or composition may begin to drift.

Therefore, to make a 20-second sequence, it may be more reliable to generate four or five separate 4–5 second clips and combine them in an editor such as CapCut.

Optimizing Negative Prompts

A short but effective negative-prompt list can help reduce unwanted defects:

“no camera shake”

“no jump cuts”

“no glitching limbs or faces”

“no flickering shadows” — especially useful when lighting appears unstable.

“no extreme facial expressions” — useful when an avatar's face becomes distorted during a shot.

“no text overlays” — useful when you do not want the AI to place unclear generated lettering on the video.

AI Video Evaluation and VBench-2.0

How can we evaluate objectively whether a generated video is excellent or poor? Researchers have developed benchmarks such as VBench for this purpose. When you generate a video, it is useful to know what constitutes a technically "bad" result rather than judging only by first impression.

From Superficial Faithfulness to Intrinsic Faithfulness

Early AI video systems were often evaluated mainly for "superficial faithfulness": whether the video looked good and appeared to match the text prompt, including factors such as frame aesthetics, temporal smoothness, and basic text alignment. But such models could still violate physical laws or human anatomy.

VBench-2.0 is presented as an attempt to evaluate "intrinsic faithfulness" in AI video. A video should not only look appealing; it should also obey the physical and logical rules of the real world. The source describes five major dimensions:

Human Fidelity: evaluates structural correctness of the human figure and temporal consistency across generated video.

Structural correctness: checks whether body parts such as hands or legs suddenly disappear or become distorted. For example, if a character turns, the body should not deform unnaturally into what the source calls a "thousand-hand yoga" effect.

Temporal consistency — identity and clothing: the source refers to ArcFace and RetinaFace for face identity and LLaVA-video-7B for clothing consistency, with the goal of checking that the same person remains identifiable and clothing colour or design does not change unexpectedly across shots.

Creativity: measures how much variety a model can produce from the same prompt. The source mentions pre-trained VGG-19 feature representations as one method of assessing diversity.

Controllability: measures how accurately the model follows instructions such as specific camera movements, including pan and zoom.

Physics: examines whether gravity, light reflection, and other physical rules are respected in the video.

Commonsense: checks whether the sequence appears logically coherent and plausible.

Figure Description 7: a graphical representation of a VBench leaderboard, rating AI models such as Runway and Sora on a normalized scale from about 0.3 to 0.8 across criteria such as human fidelity, creativity, and physics.

The source further describes VBench++ as supporting evaluation not only for text-to-video but also image-to-video models. It is described as open source, with adaptive-aspect-ratio support and additional attention to trustworthiness.

Practical Workflow for the Complete Production Process

Becoming a successful video creator requires more than knowing individual tools; the entire workflow must be organized and applied systematically. This Teach Yourself section brings the whole process together in one place.

API Integration and Advanced Parameters

Technical users can control parameters such as aspect ratio, quality, output format, camera behaviour, and clip chaining through the APIs of some advanced video tools. Names and availability differ among systems; the examples below are included to explain the concepts.

aspect_ratio: 16:9 for YouTube, 9:16 for Reels, 1:1, 4:3, 21:9 for cinematic use, or “adaptive,” allowing the model to determine the ratio from the first frame or prompt.

bitrate_mode: “standard” (CRF 18) or “high” (CRF 11). The source describes high mode as producing files roughly 3–5 times larger while retaining finer detail.

output_format: .mov with yuv444p or yuv444p10le is described as useful for high colour fidelity and repeated extension/edit chains, while .mp4 with yuv420p is presented as a common default for direct social-media upload.

camera_fixed: setting this to true biases the model toward a locked-off camera, meaning the camera remains still.

return_last_frame: setting this to true is described as returning the final frame as a watermark-free PNG that can be fed back as input for chaining, allowing a short clip to be extended through successive generations.

Final Assembly: Step by Step

A complete workflow for creating an AI video in Maithili can be organized as follows:

Story and prompt development: first choose the subject and prepare a clear English prompt using the prompt-engineering framework described earlier in the book.

Video generation: use an appropriate available text-to-video or image-to-video tool to create short clips. If a clip contains distortion, continuity errors, or unnatural movement, reject it against your quality criteria and generate a new version.

Audio production: use your own recorded Maithili voice or an appropriate text-to-speech tool. If synthetic speech is used, check pronunciation, names, pauses, emotion, and consent; clone a voice only when you have clear rights and permission.

Lip-sync and avatars: if a speaking character appears in the video, use an appropriate lip-sync or avatar tool to align the image and audio. Check mouth movement, expression, timing, and character identity separately.

Final editing (post-production): in a suitable timeline editor, combine the short clips, narration, sound effects, background music, titles, and subtitles. Before final export, check visual continuity, audio balance, and textual accuracy (Dancyger, 2026; Clark et al., 2019).

Figure Description 8: a CapCut desktop timeline containing four short video clips, an audio track generated with ElevenLabs below them, and a background-music layer beneath.

Conclusion

The study above shows that AI-based video production is becoming established as a practical creative tool. Generative models can rapidly produce many variations of visuals, motion, voice, and style, but human checking remains necessary for physical realism, character continuity, factual accuracy, cultural precision, and ethical judgment (Enochs, 2026; Anderson, 2026; Hutson and Smith, 2025).

With balanced prompt engineering, control over camera angle and movement, appropriate frame-rate choices such as 24 fps, and a sound understanding of lighting, a new user can use the methods in this Teach Yourself book to create high-quality cinematic video in Maithili or another language. These technologies offer creators wide creative freedom and control, allowing ideas to be brought vividly to a technical medium.

Selected Bibliography

This bibliography includes books only. These works are useful for durable principles of generative media, prompt writing, visual composition, editing, sound, documentary practice, post-production, and AI ethics. Tool-specific buttons, plans, prices, and versions are temporary; the book therefore places its main emphasis on lasting skills and workflows.

Anderson, Iain. 2026. AI for Creative Production: A Handbook for the Ethical Use of AI in Creating and Processing Text, Images, Video, and Audio. 1st ed. Packt Publishing.

Block, Bruce. 2021. The Visual Story: Creating the Visual Structure of Film, TV, and Digital Media. 3rd ed. Routledge.

Bowen, Christopher. 2024. Grammar of the Shot. 5th ed. Routledge.

Brown, Blain. 2022. Cinematography: Theory and Practice: For Cinematographers and Directors. 4th ed. Routledge.

Chattopadhyay, Budhaditya. 2024. Sound in Indian Film and Audiovisual Media: History, Practices and Production. 1st ed. Routledge.

Clark, Barbara, Susan Spohr, Dawn Higginbotham, and Kumari Bakhru. 2019. The Guide to Managing Postproduction for Film, TV, and Digital Distribution: Managing the Process. 3rd ed. Routledge.

Coeckelbergh, Mark. 2020. AI Ethics. The MIT Press.

Dancyger, Ken. 2026. The Technique of Film and Video Editing: History, Theory, and Practice. 7th ed. Routledge.

Dubber, Markus D., Frank Pasquale, and Sunit Das, eds. 2020. The Oxford Handbook of Ethics of AI. Oxford University Press.

Enochs, Joseph. 2026. Video Generation with AI. O’Reilly Media.

Farnell, Andy. 2010. Designing Sound. The MIT Press.

Foster, David. 2023. Generative Deep Learning. 2nd ed. O’Reilly Media.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. Deep Learning. The MIT Press.

Hutson, James, and Andrew Smith. 2025. Cinematic Algorithms: The Rise of Generative AI in Video Art and Visual Culture. CRC Press.

Khan, Hussin. 2024. Virtual Filmmaking with Unreal Engine 5: A Step-by-Step Guide to Creating a Complete Animated Short Film. 1st ed. Packt Publishing.

Liao, S. Matthew, ed. 2020. Ethics of Artificial Intelligence. Oxford University Press.

Phoenix, James, and Mike Taylor. 2024. Prompt Engineering for Generative AI. 1st ed. O’Reilly Media.

Rabiger, Michael, and Courtney Hermann. 2026. Directing the Documentary. 8th ed. Routledge.

Rose, Jay. 2015. Producing Great Sound for Film and Video: Expert Tips from Preproduction to Final Mix. 4th ed. Routledge.

Russell, Stuart, and Peter Norvig. 2020. Artificial Intelligence: A Modern Approach. 4th ed. Pearson.

Sanseviero, Omar, Pedro Cuenca, Apolinário Passos, and Jonathan Whitaker. 2024. Hands-On Generative AI with Transformers and Diffusion Models. 1st ed. O'Reilly Media.

This bibliography is intended for thematic reading. It does not treat the temporary version, price, or interface of any one commercial tool described in the book as a permanent standard.

Videha

Gajendra Thakur
Editor — Videha Maithili e-Journal
www.videha.co.in
ISSN 2229-547X